116229

STIC-Biotech/ChemLib

From: Bunner, Bridget
Sent: Monday, March 08, 2004 11:59 AM
To: STIC-Biotech/ChemLib
Subject: sequence search

Hi! I'd like to request a sequence search for case 09/555,296:

1. the amino acid sequence of SEQ ID NO: 4

URK

Thanks!

Bridget Bunner

Art Unit 1647
Rem 4C65
(571) 272-0881
mailbox 4C70

Searcher:_____
Phone:_____
Location:_____
Date Picked Up:_____
Date Completed:_____
Searcher Prep/Review:_____
Clerical:_____
Online time:_____

TYPE OF SEARCH:
NA Sequences:_____
AA Sequences:_____
Structures:_____
Bibliographic:_____
Litigation:_____
Full text:_____
Patent Family:_____
Other:_____

VENDOR/COST (where applic.)
STN:_____
DIALOG:_____
Questel/Orbit:_____
DRLink:_____
Lexis/Nexis:_____
Sequence Sys.:_____
WWW/Internet:_____
Other (specify):_____

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 9, 2004, 13:32:25 ; Search time 23 Seconds
(without alignments)
469.123 Million cell updates/sec

Title: US-09-555-296B-4
Perfect score: 1125
Sequence: 1 MKMQVVLLLTFVSAALATQA.....DRKTYDIFNEECVYNGEPWL 209

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1125	100.0	209	4	US-09-618-919A-22	Sequence 22, Appl
2	290.5	25.8	190	4	US-09-618-919A-18	Sequence 18, Appl
3	265	23.6	200	4	US-09-618-919A-20	Sequence 20, Appl
4	252	22.4	190	4	US-09-618-919A-16	Sequence 16, Appl
5	102.5	9.1	284	4	US-09-134-000C-3594	Sequence 3594, Ap
6	91	8.1	202	1	US-07-807-022A-1	Sequence 1, Appli
7	84.5	7.5	497	4	US-09-107-532A-5279	Sequence 5279, Ap
8	83	7.4	510	4	US-09-134-000C-6635	Sequence 6635, Ap
9	83	7.4	2285	4	US-09-308-375-2	Sequence 2, Appli
10	81	7.2	659	3	US-08-894-818B-1	Sequence 1, Appli
11	81	7.2	659	3	US-08-894-818B-5	Sequence 5, Appli
12	81	7.2	659	4	US-09-445-472-12	Sequence 12, Appl
13	79.5	7.1	913	1	US-08-487-890A-5	Sequence 5, Appli
14	79.5	7.1	913	2	US-08-478-435-5	Sequence 5, Appli
15	79.5	7.1	913	2	US-08-337-483-5	Sequence 5, Appli
16	79.5	7.1	913	2	US-08-478-373-5	Sequence 5, Appli
17	79.5	7.1	913	3	US-08-474-671-5	Sequence 5, Appli
18	79.5	7.1	913	3	US-08-483-577A-5	Sequence 5, Appli
19	79.5	7.1	913	3	US-08-897-438-5	Sequence 5, Appli
20	79.5	7.1	913	4	US-08-637-654-5	Sequence 5, Appli
21	79.5	7.1	913	4	US-08-649-518-5	Sequence 5, Appli
22	78	6.9	1871	2	US-08-694-869-1	Sequence 1, Appli
23	78	6.9	1871	3	US-09-349-546-1	Sequence 1, Appli
24	78	6.9	1871	4	US-09-502-831-1	Sequence 1, Appli
25	77.5	6.9	652	4	US-09-489-039A-12461	Sequence 12461, A
26	77	6.8	640	4	US-09-328-352-8058	Sequence 8058, Ap
27	77	6.8	914	1	US-08-487-890A-11	Sequence 11, Appl
28	77	6.8	914	2	US-08-478-435-11	Sequence 11, Appl
29	77	6.8	914	2	US-08-337-483-11	Sequence 11, Appl
30	77	6.8	914	2	US-08-478-373-11	Sequence 11, Appl
31	77	6.8	914	3	US-08-474-671-11	Sequence 11, Appl
32	77	6.8	914	3	US-08-483-577A-11	Sequence 11, Appl
33	77	6.8	914	3	US-08-897-438-11	Sequence 11, Appl
34	77	6.8	914	4	US-08-637-654-11	Sequence 11, Appl
35	77	6.8	914	4	US-08-649-518-11	Sequence 11, Appl
36	76.5	6.8	406	1	US-08-434-881-2	Sequence 2, Appli
37	76.5	6.8	406	3	US-08-977-771-2	Sequence 2, Appli
38	76.5	6.8	406	3	US-09-361-773-2	Sequence 2, Appli
39	76	6.8	416	4	US-09-540-236-2388	Sequence 2388, Ap
40	75.5	6.7	391	4	US-09-107-532A-4419	Sequence 4419, Ap
41	75	6.7	179	4	US-09-328-352-6376	Sequence 6376, Ap
42	75	6.7	246	4	US-09-107-532A-5142	Sequence 5142, Ap
43	75	6.7	912	1	US-08-487-890A-7	Sequence 7, Appli
44	75	6.7	912	1	US-08-487-890A-9	Sequence 9, Appli
45	75	6.7	912	2	US-08-478-435-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-09-618-919A-22
; Sequence 22, Application US/09618919A
; Patent No. 6617312
; GENERAL INFORMATION:
; APPLICANT: Paesen, Guido Christiaan
; APPLICANT: Nuttall, Patricia Anne
; TITLE OF INVENTION: Vasoactive Amine Binding Molecules
; FILE REFERENCE: 2369-1-001CON
; CURRENT APPLICATION NUMBER: US/09/618,919A
; CURRENT FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: US 09/180,733
; PRIOR FILING DATE: 1998-11-13
; PRIOR APPLICATION NUMBER: GB 9610484.0
; PRIOR FILING DATE: 1996-05-18
; PRIOR APPLICATION NUMBER: GB 9707844.8
; PRIOR FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Dermacentor reticulatus
US-09-618-919A-22

Query Match 100.0%; Score 1125; DB 4; Length 209;
Best Local Similarity 100.0%; Pred. No. 2.8e-114;
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MKMQVVLLLTFVSAALATQAETTSAKAGENPLWAHEELLGKYQDAWKSIDQGVSVTYVLA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKMQVVLLLTFVSAALATQAETTSAKAGENPLWAHEELLGKYQDAWKSIDQGVSVTYVLA 60

Qy         61 KTTYENDTGSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASSPIKYYNVTETVKAVFQY 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 KTTYENDTGSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASSPIKYYNVTETVKAVFQY 120

Qy        121 GYKNIRNAIEYQVGGGLNITDTLIFTDGELCDVFYVPNADQGCELWVKKSHYKHVPDYCT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GYKNIRNAIEYQVGGGLNITDTLIFTDGELCDVFYVPNADQGCELWVKKSHYKHVPDYCT 180

Qy        181 FVFNVFCAKDRKTYDIFNEECVYNGEPWL 209
              |||||||||||||||||||||||||||||
Db        181 FVFNVFCAKDRKTYDIFNEECVYNGEPWL 209
```

RESULT 2
US-09-618-919A-18
; Sequence 18, Application US/09618919A
; Patent No. 6617312

```
; GENERAL INFORMATION:
;  APPLICANT: Paesen, Guido Christiaan
;  APPLICANT:  Nuttall, Patricia Anne
;  TITLE OF INVENTION: Vasoactive Amine Binding Molecules
;  FILE REFERENCE: 2369-1-001CON
;  CURRENT APPLICATION NUMBER: US/09/618,919A
;  CURRENT FILING DATE:  2000-07-18
;  PRIOR APPLICATION NUMBER: US 09/180,733
;  PRIOR FILING DATE: 1998-11-13
;  PRIOR APPLICATION NUMBER: GB 9610484.0
;  PRIOR FILING DATE: 1996-05-18
;  PRIOR APPLICATION NUMBER: GB 9707844.8
;  PRIOR FILING DATE: 1997-04-18
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 18
;   LENGTH: 190
;   TYPE: PRT
;   ORGANISM: Rhipicephalus appendiculatus
US-09-618-919A-18

  Query Match             25.8%;  Score 290.5;  DB 4;  Length 190;
  Best Local Similarity   36.0%;  Pred. No. 1.3e-23;
  Matches   72;  Conservative   32;  Mismatches   79;  Indels   17;  Gaps    9;

Qy          7 LLLTFVSAALATQAETTSAKAGENPLWAHEELLGKYQDAWKSIDQGVSVTYVLAKTTYEN 66
              ||:  ::  ||       |   | | || |   | :||||||:   |   | : | ||:|
Db          3 LLILSLALVLA-----LSQVKGNQPDWADEAANGAHQDAWKSLKADVENVYYMVKATYKN 57

Qy         67 DTGSWGSQFKCLQVQEIERKEEDYTVTSVFTF-RNASSPIKYYNVTETVKAVFQYGYKNI 125
              |   ||: | |: |   :  |:: :: : | |  || : :::   || | ||  ||| |
Db         58 DP-VWGNDFTCVGVMANDVNEDEKSIQAEFLFMNNADTNMQF--ATEKVTAVKMYGY-NR 113

Qy        126 RNAIEYQVGGGLNITDTLIFTDGELCDVFYVPNAD---QGCELWVKKSHYKHVPDYCTFV 182
               ||  |:   |   || : ::| : ||| |||  |   :| |||   : | ::|  |
Db        114 ENAFRYETEDGQVFTDVIAYSD-DNCDVIYVPGTDGNEEGYELWT--TDYDNIPANCLNK 170

Qy        183 FNVFCAKDRKTYDIFNEECV 202
              || : |  |:| |:|   |:
Db        171 FNEY-AVGRETRDVFTSACL 189


RESULT 3
US-09-618-919A-20
; Sequence 20, Application US/09618919A
; Patent No. 6617312
; GENERAL INFORMATION:
;  APPLICANT: Paesen, Guido Christiaan
;  APPLICANT:  Nuttall, Patricia Anne
;  TITLE OF INVENTION: Vasoactive Amine Binding Molecules
;  FILE REFERENCE: 2369-1-001CON
;  CURRENT APPLICATION NUMBER: US/09/618,919A
;  CURRENT FILING DATE:  2000-07-18
;  PRIOR APPLICATION NUMBER: US 09/180,733
;  PRIOR FILING DATE: 1998-11-13
;  PRIOR APPLICATION NUMBER: GB 9610484.0
;  PRIOR FILING DATE: 1996-05-18
;  PRIOR APPLICATION NUMBER: GB 9707844.8
;  PRIOR FILING DATE: 1997-04-18
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 20
;   LENGTH: 200
;   TYPE: PRT
;   ORGANISM: Rhipicephalus appendiculatus
US-09-618-919A-20

  Query Match             23.6%;  Score 265;  DB 4;  Length 200;
  Best Local Similarity   34.5%;  Pred. No. 8.5e-21;
  Matches   69;  Conservative   34;  Mismatches   71;  Indels   26;  Gaps   10;

Qy          3 MQVVLLLTFVSAALATQAETTSAKAGENPLWAHEELLGKYQDAWKSIDQGVSVTYVLAKT 62
              |:|:||:  : |||   |:        || ||:|  || |||||||: |  :  | ||:

Db          1 MKVLLLV--LGAALCQNADA-------NPTWANEAKLGSYQDAWKSLQQDQNKRYYLAQA 51

Qy         63 TYENDTGSWGSQFKCLQV--QEIERKEEDYTVTSVFTFRNAS-SPIKYYNVTETVKAVFQ 119
              |   | | || :| |: |  ::| :|: : |:      ::|   : :|  : | ||   :
Db         52 TQTTD-GVWGEEFTCVSVTAEKIGKKKLNATI----LYKNKHLTDLKESHETITVWKAYD 106

Qy        120 YGYKNIRNAIEYQVGGGLNIT--DTLIFTDGELCDVFYVP---NADQG-CELWVKKSHYK 173
              |       | |:|:  |    |  | :|:| : ||| :||    :|:|  |||| :
Db        107 Y---TTENGIKYETQGTRTQTFEDVFVFSDYKNCDVIFVPKERGSDEGDYELWVSEDKID 163

Qy        174 HVPDYCTFVFNVFCAKDRKT 193
               :|| | |    | :  ||
Db        164 KIPDCCKFTMAYFAQQQEKT 183


RESULT 4
US-09-618-919A-16
; Sequence 16, Application US/09618919A
; Patent No. 6617312
; GENERAL INFORMATION:
;  APPLICANT: Paesen, Guido Christiaan
;  APPLICANT:  Nuttall, Patricia Anne
;  TITLE OF INVENTION: Vasoactive Amine Binding Molecules
;  FILE REFERENCE: 2369-1-001CON
;  CURRENT APPLICATION NUMBER: US/09/618,919A
;  CURRENT FILING DATE:  2000-07-18
;  PRIOR APPLICATION NUMBER: US 09/180,733
;  PRIOR FILING DATE: 1998-11-13
;  PRIOR APPLICATION NUMBER: GB 9610484.0
;  PRIOR FILING DATE: 1996-05-18
;  PRIOR APPLICATION NUMBER: GB 9707844.8
;  PRIOR FILING DATE: 1997-04-18
;  NUMBER OF SEQ ID NOS: 24
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16
;   LENGTH: 190
;   TYPE: PRT
;   ORGANISM: Rhipicephalus appendiculatus
US-09-618-919A-16

  Query Match             22.4%;  Score 252;  DB 4;  Length 190;
  Best Local Similarity   33.5%;  Pred. No. 2e-19;
  Matches   68;  Conservative   28;  Mismatches   89;  Indels   18;  Gaps    8;

Qy          3 MQVVLLLTFVSAALATQAETTSAKAGENPLWAHEELLGKYQDAWKSIDQGVSVTYVLAKT 62
              |:::| | || |    :|       : |:|| |   |::|||||: : : |   | | |
Db          1 MKLLLSLAFVLALSQVKA--------DKPVWADEAANGEHQDAWKHLQKLVEENYDLIKA 52

Qy         63 TYENDTGSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASSPIKYYNVTETVKAVFQYGY 122
              ||:||   ||: | |:          |::  | : | | | :  : | :  |       |||
Db         53 TYKNDP-VWGNDFTCVGTAAQNLNEDEKNVEAWFMFMNNADTV-YQHTFEKATPDKMYGY 110

Qy        123 KNIRNAIEYQVGGGLNITDTLIFTDGELCDVFYVPNAD---QGCELWVKKSHYKHVPDYC 179
               |  ||| ||   |  :|| | |:| : | | |     |  | |||   : |  ||  |
Db        111 -NKENAITYQTEDGQVLTDVLAFSD-DNCYVIYALGPDGSGAGYELWA--TDYTDVPASC 166

Qy        180 TFVFNVFCAKDRKTYDIFNEECV 202
                || : |      |::  :|:
Db        167 LEKFNEYAA-GLPVRDVYTSDCL 188


RESULT 5
US-09-134-000C-3594
; Sequence 3594, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;  TITLE OF INVENTION:  ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 032796-032
;  CURRENT APPLICATION NUMBER: US/09/134,000C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/055,778
```

```
;  PRIOR FILING DATE: 1997-08-15
;  NUMBER OF SEQ ID NOS: 6812
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3594
;   LENGTH: 284
;   TYPE: PRT
;   ORGANISM: Enterococcus faecalis
US-09-134-000C-3594

  Query Match             9.1%;  Score 102.5;  DB 4;  Length 284;
  Best Local Similarity   26.9%;  Pred. No. 0.0066;
  Matches   60;  Conservative   26;  Mismatches   72;  Indels   65;  Gaps   15;

Qy         16 LATQAETTSA-----KAGENPLWAHEELLGKYQDAWKSI-DQGVSVTYVLAKTTY----- 64
Db         68 LVGDKETGSSYGFAVKKGQNP-----ELIKKFNAGLKNLKDNG---TYDKILNNYLATGD 119

Qy         65 ENDTGSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASSPIKYYNV--------TETVKA 116
Db        120 ETNTQDAGEQMK-----KITPKKEKYVIASDSTF----APFEFQNAQGDYVGIDVDLVKR 170

Qy        117 V---------FQY-GYKNIRNAIEY-QVGG---GLNITDTLIFTDGELCDVFYVPNADQG 162
Db        171 AAELQGFTVEFKFIGFSSAVQAVESGQADGMVAGMTITD-----DRKKAFDFSVPYFDSG 225

Qy        163 CELWVKKSH--YKHVPDYCTFVFNV--------FCAKDRKTYD 195
Db        226 IQIAVKKGNDKIKSYDDLKGKKVGVKIGTESADFLEKNKKKYD 268


RESULT 6
US-07-807-022A-1
; Sequence 1, Application US/07807022A
; Patent No. 5321010
;  GENERAL INFORMATION:
;    APPLICANT:  Connolly, Thomas M.
;    APPLICANT:  Neeper, Michael
;    APPLICANT:  Waxman, Lloyd
;    TITLE OF INVENTION:  Protein for Inhibiting Adhesion of
;    TITLE OF INVENTION:  Platelets to Collagen
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merck & Co., Inc.
;      STREET:  P.O. Box 2000
;      CITY:  Rahway
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  08840
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/807,022A
;      FILING DATE:  19911210
;      CLASSIFICATION:  514
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Parr, Richard S.
;      REGISTRATION NUMBER:  32,586
;      REFERENCE/DOCKET NUMBER:  18143
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (908) 594-4958
;      TELEFAX:  908-594-4720
;      TELEX:  138825
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  202 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  unknown
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  protein
US-07-807-022A-1

  Query Match             8.1%;  Score 91;  DB 1;  Length 202;
  Best Local Similarity   20.8%;  Pred. No. 0.072;
  Matches   46;  Conservative   33;  Mismatches   82;  Indels   60;  Gaps   11;

Qy          2 KMQVVLLLTFVSAALATQAETTSAKAGENPLWAHEELLGKYQDAWKSIDQGVSVTYVLAK 61
Db          1 KLAILAVCVFISTVSSVPTSDTSEE---------EDLNTKNRDIWKTYSK--SKVFWLLR 49

Qy         62 TTYENDTGSWGSQFKCL-----------QVQEI----ERKEEDYTVTSVF-TFRNASSP 104
Db         50 RTYWVD----GKRSLCRYGTVLKRDKANHRIEQIMGSYDESQETYTKTKVYITTRLGRSG 105

Qy        105 IKYYNVTETVKAVFQYGYKNIRNAIEYQVGGGLNITDTLIFTDGELCDVFYV-------- 156
Db        106 DRNH------MGVSLQGYNH--TGIEYK----------MIYDDDQGCAILKVTKDNRNQP 147

Qy        157 PNADQGCELWVKKSHYKHVPDY--CTFVFNVFCAKDRKTYD 195
Db        148 QNLKNACEMWATANDANSVNSIAACEVVYQRRCNPNNSVDD 188


RESULT 7
US-09-107-532A-5279
; Sequence 5279, Application US/09107532A
; Patent No. 6583275
;  GENERAL INFORMATION:
;       APPLICANT: Lynn A Doucette-Stamm  and David Bush
;       TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;                           ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
;       NUMBER OF SEQUENCES: 7310
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: GENOME THERAPEUTICS CORPORATION
;            STREET: 100 Beaver Street
;            CITY: Waltham
;            STATE: Massachusetts
;            COUNTRY: USA
;            ZIP: 02354
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: CD/ROM ISO9660
;            COMPUTER: PC
;            OPERATING SYSTEM: <Unknown>
;            SOFTWARE: ASCII
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/107,532A
;            FILING DATE: 30-Jun-1998
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 60/085,598
;            FILING DATE: 14 May 1998
;            APPLICATION NUMBER: 60/051571
;            FILING DATE: July 2, 1997
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Ariniello, Pamela Deneke
;            REGISTRATION NUMBER: 40,489
;            REFERENCE/DOCKET NUMBER: GTC-012
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (781)893-5007
;            TELEFAX: (781)893-8277
;  INFORMATION FOR SEQ ID NO: 5279:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 497 amino acids
;            TYPE: amino acid
;            TOPOLOGY: linear
;       MOLECULE TYPE: protein
;       HYPOTHETICAL: YES
;       ORIGINAL SOURCE:
;            ORGANISM: Enterococcus faecium
;       FEATURE:
;            NAME/KEY:  misc_feature
;            LOCATION: (B) LOCATION 1...497
;       SEQUENCE DESCRIPTION: SEQ ID NO: 5279:
US-09-107-532A-5279
```

```
  Query Match             7.5%;  Score 84.5;  DB 4;  Length 497;
  Best Local Similarity   22.6%;  Pred. No. 1.4;
  Matches   47;  Conservative   29;  Mismatches   85;  Indels   47;  Gaps    9;

Qy         26 KAGENPLWAHEELLGKYQDAWKSIDQGVSVTYVLAKTTYEND--TGSWGSQFKCLQVQEI 83
              |  |  ||        : : ||   ::|:||: |    |:::|     |  |:  :  |::
Db         15 KFPEGFLWGGATAANQCEGAWDVDEKGLSVSDVY---TFDSDLPKEKWTDQWHMMTHQQV 71

Qy         84 ERKEEDYTVTSVFTFRNASSPIKYYNVTETVKAVFQYGYKNIRNAIEYQVGGGLNITDTL 143
              :  ::     | : :       |::  | ::   : |:|  | :| :          |
Db         72 KAAQDP---NSKYYYPKRHGNDFYHHFKEDIRLFAEMGFKCYRMSIAW----------TR 118

Qy        144 IF------TDGELCDVFYVPNADQ----GCELWVKKSHY----------------KHVPD 177
              ||      |  |    ||    |:    | |  |  |||                | :
Db        119 IFPHGDEETPNEAGLFFYDQVFDECLKYGIEPVVSLSHYEMPLYLVTEYGGWPNRKLIQF 178

Qy        178 YCTFVFNVFCAKDRKT--YDIFNE-ECV 202
              |  |   ||      |   :  ||| ||
Db        179 YVRFAETVFKRYKNKVKYWMTFNEINCV 206


RESULT 8
US-09-134-000C-6635
; Sequence 6635, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
;  APPLICANT: Lynn Doucette-Stamm et al
;  TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
;  TITLE OF INVENTION:  ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
;  FILE REFERENCE: 032796-032
;  CURRENT APPLICATION NUMBER: US/09/134,000C
;  CURRENT FILING DATE:  1998-08-13
;  PRIOR APPLICATION NUMBER: US 60/055,778
;  PRIOR FILING DATE: 1997-08-15
;  NUMBER OF SEQ ID NOS: 6812
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6635
;   LENGTH: 510
;   TYPE: PRT
;   ORGANISM: Enterococcus faecalis
US-09-134-000C-6635

  Query Match             7.4%;  Score 83;  DB 4;  Length 510;
  Best Local Similarity   20.2%;  Pred. No. 2.1;
  Matches   49;  Conservative   42;  Mismatches   77;  Indels   74;  Gaps   13;

Qy          5 VVLLLTFVSAALATQAETTSAKAGENPL--WAHE-ELLGK---YQDAWKSIDQGVSVTYV 58
              | :: : : ||:    ::|| :  |  :  : || :::     | |     |:|     |
Db          9 VKIIASALKAAIDKDRKSTSKREAETGIRYYNHENDIMNNRIFYVD-----DEG-----V 58

Qy         59 LAKTTYENDTGSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASSPIK-----YYN---- 109
              | :  | ::       |       ||  ::  | :::   :   :  :|     |||
Db         59 LREDKYASNVRIPHGFF-----PEIVDQKTQYLLSNPVEYETENEELKEYLAEYYNSEFQ 113

Qy        110 --VTETVKAVFQYGYK------NIRNAIEYQVGGGLNITD----------------TLIF 145
                : | |:   | |::      |  : : :||   ||:                   | |
Db        114 VVLQELVEGSSQKGFEYVYARTNAEDRLCFQVADSLNVFGVYNEYNELQRICRHYITEIE 173

Qy        146 TDGELCDVFYVPNADQGCELWVKKSHYKHVPDYCTFVFNVFCAKDRKTYDIFNEECVYNG 205
               |||  |: :        |:|  :: |             | |:| | |::   | : |
Db        174 KDGETVDIHH-------AEVWTDQNVY-----------FFVAEDNKDYELDEAEPI-NP 213

Qy        206 EP 207
               |
Db        214 RP 215


RESULT 9
US-09-308-375-2
; Sequence 2, Application US/09308375
; Patent No. 6300117
; GENERAL INFORMATION:
;  APPLICANT: Genencor International, Inc.
;  TITLE OF INVENTION: Proteases From Gram-Positive Organisms
;  FILE REFERENCE: GC394-PCT
;  CURRENT APPLICATION NUMBER: US/09/308,375
;  CURRENT FILING DATE: 1999-05-14
;  EARLIER APPLICATION NUMBER: EP9719636.4
;  EARLIER FILING DATE: 1997-09-15
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 2285
;   TYPE: PRT
;   ORGANISM: Bacillius subtilis
US-09-308-375-2

  Query Match             7.4%;  Score 83;  DB 4;  Length 2285;
  Best Local Similarity   23.3%;  Pred. No. 19;
  Matches   51;  Conservative   27;  Mismatches   79;  Indels   62;  Gaps   10;

Qy          6 VLLLTFVSAALATQAETTSAKAGENPLWAHEELLGKYQDAWKSID--QGVSVTYVLAKTT 63
              :|: | |  | ||                  || | |:  : :| |: |  :    | | | ||
Db        743 LLVSTLVGGAFAALG-----------WALESLISSFAEAKKAKDDFEQSQQTNVEAITT 790

Qy         64 YENDTGSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASSPIKYYNVTETVKAVFQ---Y 120
               :: |     |:|     |::: :|  ::||          :|  ||: :   |
Db        791 NKDSTDKLIQQYK-----ELQKVKESRSLTS-------DEEQEYLQVTQQLAQTFPALVK 838

Qy        121 GY-----------KNIRNAIEYQVGGGLNITDTLIFTDGELCDVFYVPNADQGCELWVKK 169
              ||           | :  |||       |  : |     |  |        |  |  :||
Db        839 GYDSQGNAILKTNKELEKAIE-------NTKEYLALKKQETRDSAKKTFEDASKE--IKK 889

Qy        170 S-----HYKHVPDYCTFVFNVFCAKDRKTYDIFNEECVY 203
              |     || : ||          | |  :|:  ::  |
Db        890 SKDELKQYKQIADYND--------KGRPKWDLIADDDDY 920


RESULT 10
US-08-894-818B-1
; Sequence 1, Application US/08894818B
; Patent No. 6261822
;  GENERAL INFORMATION:
;    APPLICANT:  TAKAKURA, Hikaru
;    APPLICANT:  MORISHITA, Mio
;    APPLICANT:  YAMAMOTO, Katsuhiko
;    APPLICANT:  MITTA, Masanori
;    APPLICANT:  ASADA, Kiyozo
;    APPLICANT:  TSUNASAWA, Susumu
;    APPLICANT:  KATO, Ikunoshin
;    TITLE OF INVENTION:  HYPERTHERMOSTABLE PROTEASE GENES
;    NUMBER OF SEQUENCES:  42
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Browdy and Neimark
;      STREET:  419 Seventh Street N.W., Ste. 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  United States of America
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/894,818B
;      FILING DATE:  20-MAY-1998
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/JP96/03253
;      FILING DATE:  07-NOV-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 323285/1995
;      FILING DATE:  12-DEC-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Browdy, Roger L.
```

```
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  TAKAKURA=1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 628-5197
;      TELEFAX:  (202) 737-3528
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  659 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-894-818B-1

  Query Match             7.2%;  Score 81;  DB 3;  Length 659;
  Best Local Similarity   25.9%;  Pred. No. 5;
  Matches   53;  Conservative   20;  Mismatches   62;  Indels   70;  Gaps   13;

Qy         41 KYQDAWK------SIDQGVSV-TYVLAKTTYENDTGSW--GSQFKCLQVQEIERKEEDYT 91
Db        436 KYDDYAKLTFTGSVADKGSATHTFDVSGATFVTATLYWDTGSSDIDLYLYDPNGNEVDYS 495

Qy         92 VTSVFTFRNASSPIKYYNVTE---TVKAVFQYGYKNIRNAIEYQV----------GGG- 136
Db        496 YTAYYGFEK----VGYYNPTAGTWTVKVV---SYKGAAN---YQVDVVSDGSLSQSGGGN 545

Qy        137 ---------LNITDTLIFTDGELCDVFYVPNADQGCELWVKKSHYKHVPDYCTFVFNVFC 187
Db        546 PNPNPNPNPTPTTDTQTFT-GSVND--YWDTSD-----------------TFTMNVNS 583

Qy        188 AKDRKTYDI-----FNEECVYNGEP 207
Db        584 GATKITGDLTFDTSYNDLDLYLYDP 608


RESULT 11
US-08-894-818B-5
; Sequence 5, Application US/08894818B
; Patent No. 6261822
;  GENERAL INFORMATION:
;    APPLICANT:  TAKAKURA, Hikaru
;    APPLICANT:  MORISHITA, Mio
;    APPLICANT:  YAMAMOTO, Katsuhiko
;    APPLICANT:  MITTA, Masanori
;    APPLICANT:  ASADA, Kiyozo
;    APPLICANT:  TSUNASAWA, Susumu
;    APPLICANT:  KATO, Ikunoshin
;    TITLE OF INVENTION:  HYPERTHERMOSTABLE PROTEASE GENES
;    NUMBER OF SEQUENCES:  42
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Browdy and Neimark
;      STREET:  419 Seventh Street N.W., Ste. 300
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  United States of America
;      ZIP:  20004
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/894,818B
;      FILING DATE:  20-MAY-1998
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/JP96/03253
;      FILING DATE:  07-NOV-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 323285/1995
;      FILING DATE:  12-DEC-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Browdy, Roger L.
```

```
;      REGISTRATION NUMBER:  25,618
;      REFERENCE/DOCKET NUMBER:  TAKAKURA=1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202) 628-5197
;      TELEFAX:  (202) 737-3528
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  659 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-894-818B-5

  Query Match             7.2%;  Score 81;  DB 3;  Length 659;
  Best Local Similarity   25.9%;  Pred. No. 5;
  Matches   53;  Conservative   20;  Mismatches   62;  Indels   70;  Gaps   13;

Qy         41 KYQDAWK------SIDQGVSV-TYVLAKTTYENDTGSW--GSQFKCLQVQEIERKEEDYT 91
Db        436 KYDDYAKLTFTGSVADKGSATHTFDVSGATFVTATLYWDTGSSDIDLYLYDPNGNEVDYS 495

Qy         92 VTSVFTFRNASSPIKYYNVTE---TVKAVFQYGYKNIRNAIEYQV----------GGG- 136
Db        496 YTAYYGFEK----VGYYNPTAGTWTVKVV---SYKGAAN---YQVDVVSDGSLSQSGGGN 545

Qy        137 ---------LNITDTLIFTDGELCDVFYVPNADQGCELWVKKSHYKHVPDYCTFVFNVFC 187
Db        546 PNPNPNPNPTPTTDTQTFT-GSVND--YWDTSD-----------------TFTMNVNS 583

Qy        188 AKDRKTYDI-----FNEECVYNGEP 207
Db        584 GATKITGDLTFDTSYNDLDLYLYDP 608


RESULT 12
US-09-445-472-12
; Sequence 12, Application US/09445472
; Patent No. 6358726
; GENERAL INFORMATION:
;  APPLICANT: TAKAKURA, Hikaru
;  APPLICANT:  MORISHITA, Mio
;  APPLICANT:  SHIMOJO, Tomoko
;  APPLICANT:  ASADA, Kiyozo
;  APPLICANT:  KATO, Ikunoshin
;  TITLE OF INVENTION: SYSTEM FOR EXPRPESSING HYPERTHERMOSTABLE
;  FILE REFERENCE: TAKAKURA=6
;  CURRENT APPLICATION NUMBER: US/09/445,472
;  CURRENT FILING DATE:  1999-12-06
;  PRIOR APPLICATION NUMBER: 151969/1997
;  PRIOR FILING DATE: 1997-06-10
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
;   LENGTH: 659
;   TYPE: PRT
;   ORGANISM: Thermococcus celer
US-09-445-472-12

  Query Match             7.2%;  Score 81;  DB 4;  Length 659;
  Best Local Similarity   25.9%;  Pred. No. 5;
  Matches   53;  Conservative   20;  Mismatches   62;  Indels   70;  Gaps   13;

Qy         41 KYQDAWK------SIDQGVSV-TYVLAKTTYENDTGSW--GSQFKCLQVQEIERKEEDYT 91
Db        436 KYDDYAKLTFTGSVADKGSATHTFDVSGATFVTATLYWDTGSSDIDLYLYDPNGNEVDYS 495

Qy         92 VTSVFTFRNASSPIKYYNVTE---TVKAVFQYGYKNIRNAIEYQV----------GGG- 136
Db        496 YTAYYGFEK----VGYYNPTAGTWTVKVV---SYKGAAN---YQVDVVSDGSLSQSGGGN 545

Qy        137 ---------LNITDTLIFTDGELCDVFYVPNADQGCELWVKKSHYKHVPDYCTFVFNVFC 187
```

```
Db        546 PNPNPNPNPTPTTDTQTFT-GSVND--YWDTSD-----------------TFTMNVNS 583

Qy        188 AKDRKTYDI-----FNEECVYNGEP 207
                  : | |:     :|:  :|  :|
Db        584 GATKITGDLTFDTSYNDLDLYLYDP 608


RESULT 13
US-08-487-890A-5
; Sequence 5, Application US/08487890A
; Patent No. 5708149
;  GENERAL INFORMATION:
;    APPLICANT:  Loosmore, Sheena
;    APPLICANT:  Harkness, Robin
;    APPLICANT:  Schryvers, Anthony
;    APPLICANT:  Chong, Pele
;    APPLICANT:  Gray-Owen, Scott
;    APPLICANT:  Yang, Yan-Ping
;    APPLICANT:  Murdin, Andrew
;    APPLICANT:  Klein, Michel
;    TITLE OF INVENTION:  Transferrin Receptor Genes
;    NUMBER OF SEQUENCES:  147
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sim & McBurney
;      STREET:  6th Floor, 330 Unviersity Avenue
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5G 1R7
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/487,890A
;      FILING DATE:  07-JUN-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/175,116
;      FILING DATE:  29-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/148,968
;      FILING DATE:  08-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stewart, Michael I
;      REGISTRATION NUMBER:  24,973
;      REFERENCE/DOCKET NUMBER:  1038-466 MIS:jb
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 595-1155
;      TELEFAX:  (416) 595-1163
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  913 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-487-890A-5

  Query Match             7.1%;  Score 79.5;  DB 1;  Length 913;
  Best Local Similarity   21.9%;  Pred. No. 12;
  Matches   48;  Conservative   32;  Mismatches   60;  Indels   79;  Gaps   13;

Qy         11 FVSAALATQAETTSAKAGENPLWAHEELLGKYQD-AWKS-----IDQGVSVTYVLAKTTY 64
              :|    |  : : :  || |: :      :||::: :| :       : : ::| |: | :
Db        593 YVDLGLGMRYDVSRTKANESTI-----SVGKFKNFSWNTGIVIKPTEWLDLSYRLS-TGF 646

Qy         65 EN--------------DTGSWGSQFKCLQVQEIERKEE-------DYTVTSVFTFRNASS 103
              |              || :  :||      | | :|       |:     :  | ||
Db        647 RNPSFAEMYGWRYGGKDTDVYIGKFK----PETSRNQEFGLALKGDFGNIEISHFSNA-- 700

Qy        104 PIKYYNVTETVKAVFQ--------YGYKNIRNAIEYQVGGGLNITDTLIFTDGELCDVFY 155
                | |:    : : :        ||| | :||       |:|||  | |
```

```
Db        701 ---YRNLIAFAEELSKNGTTGKGNYGYHNAQNAKLV----GVNITAQLDFNG-------- 745

Qy        156 VPNADQGCELWVKKSHYKHVPD--YCTFVFNVFCAKDRK 192
                      ||      | :|   | || :|    ||:|
Db        746 ---------LW------KRIPYGWYATFAYNRVKVKDQK 769


RESULT 14
US-08-478-435-5
; Sequence 5, Application US/08478435
; Patent No. 5922323
;  GENERAL INFORMATION:
;    APPLICANT:  Loosmore, Sheena
;    APPLICANT:  Harkness, Robin
;    APPLICANT:  Schryvers, Anthony
;    APPLICANT:  Chong, Pele
;    APPLICANT:  Gray-Owen, Scott
;    APPLICANT:  Yang, Yan-Ping
;    APPLICANT:  Murdin, Andrew
;    APPLICANT:  Klein, Michel
;    TITLE OF INVENTION:  Transferrin Receptor Genes
;    NUMBER OF SEQUENCES:  147
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sim & McBurney
;      STREET:  Suite 701, 330 University Avenue
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5G 1R7
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/478,435
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/337,483
;      FILING DATE:  08-NOV-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/175,116
;      FILING DATE:  29-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/148,968
;      FILING DATE:  08-NOV-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stewart, Michael I
;      REGISTRATION NUMBER:  24,973
;      REFERENCE/DOCKET NUMBER:  1038-462 MIS:vg
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 595-1155
;      TELEFAX:  (416) 595-1163
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  913 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-478-435-5

  Query Match             7.1%;  Score 79.5;  DB 2;  Length 913;
  Best Local Similarity   21.9%;  Pred. No. 12;
  Matches   48;  Conservative   32;  Mismatches   60;  Indels   79;  Gaps   13;

Qy         11 FVSAALATQAETTSAKAGENPLWAHEELLGKYQD-AWKS-----IDQGVSVTYVLAKTTY 64
              :|    |  : : :  || |: :      :||::: :| :       : : ::| |: | :
Db        593 YVDLGLGMRYDVSRTKANESTI-----SVGKFKNFSWNTGIVIKPTEWLDLSYRLS-TGF 646

Qy         65 EN--------------DTGSWGSQFKCLQVQEIERKEE-------DYTVTSVFTFRNASS 103
              |              || :  :||      | | :|       |:     :  | ||
```

```
Db        647 RNPSFAEMYGWRYGGKDTDVYIGKFK----PETSRNQEFGLALKGDFGNIEISHFSNA-- 700

Qy        104 PIKYYNVTETVKAVFQ--------YGYKNIRNAIEYQVGGGLNITDTLIFTDGELCDVFY 155
                  | |:    : : :         ||| | :||        |:||| | |
Db        701 ---YRNLIAFAEELSKNGTTGKGNYGYHNAQNAKLV----GVNITAQLDFNG-------- 745

Qy        156 VPNADQGCELWVKKSHYKHVPD--YCTFVFNVFCAKDRK 192
                        ||       | :|   | || :|    ||:|
Db        746 ---------LW------KRIPYGWYATFAYNRVKVKDQK 769


RESULT 15
US-08-337-483-5
; Sequence 5, Application US/08337483
; Patent No. 5922562
;  GENERAL INFORMATION:
;    APPLICANT:  Loosmore, Sheena
;    APPLICANT:  Harkness, Robin
;    APPLICANT:  Schryvers, Anthony
;    APPLICANT:  Chong, Pele
;    APPLICANT:  Gray-Owen, Scott
;    APPLICANT:  Yang, Yan-Ping
;    APPLICANT:  Murdin, Andrew
;    APPLICANT:  Klein, Michel
;    TITLE OF INVENTION:  Transferrin Receptor Genes
;    NUMBER OF SEQUENCES:  147
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Sim & McBurney
;      STREET:  Suite 701, 330 Unviersity Avenue
;      CITY:  Toronto
;      STATE:  Ontario
;      COUNTRY:  Canada
;      ZIP:  M5G 1R7
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/337,483
;      FILING DATE:  08-NOV-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stewart, Michael I
;      REGISTRATION NUMBER:  24,973
;      REFERENCE/DOCKET NUMBER:  1038-410 MIS:jb
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (416) 595-1155
;      TELEFAX:  (416) 595-1163
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  913 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-337-483-5

  Query Match             7.1%;  Score 79.5;  DB 2;  Length 913;
  Best Local Similarity   21.9%;  Pred. No. 12;
  Matches   48;  Conservative   32;  Mismatches   60;  Indels   79;  Gaps   13;

Qy         11 FVSAALATQAETTSAKAGENPLWAHEELLGKYQD-AWKS-----IDQGVSVTYVLAKTTY 64
              :|    |  : : :  || |: :      :||::: :| :       : : ::| |: | :
Db        593 YVDLGLGMRYDVSRTKANESTI-----SVGKFKNFSWNTGIVIKPTEWLDLSYRLS-TGF 646

Qy         65 EN--------------DTGSWGSQFKCLQVQEIERKEE-------DYTVTSVFTFRNASS 103
                |              ||  :  :||      |  | :|       |:    :  | ||
Db        647 RNPSFAEMYGWRYGGKDTDVYIGKFK----PETSRNQEFGLALKGDFGNIEISHFSNA-- 700

Qy        104 PIKYYNVTETVKAVFQ--------YGYKNIRNAIEYQVGGGLNITDTLIFTDGELCDVFY 155
                  | |:    : : :         ||| | :||        |:||| | |
Db        701 ---YRNLIAFAEELSKNGTTGKGNYGYHNAQNAKLV----GVNITAQLDFNG-------- 745

Qy        156 VPNADQGCELWVKKSHYKHVPD--YCTFVFNVFCAKDRK 192
                        ||       | :|   | || :|    ||:|
Db        746 ---------LW------KRIPYGWYATFAYNRVKVKDQK 769


Search completed: March  9, 2004, 13:36:04
Job time : 25 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 9, 2004, 13:33:35 ; Search time 33 Seconds
(without alignments)
1337.304 Million cell updates/sec

Title: US-09-555-296B-4
Perfect score: 1125
Sequence: 1 MKMQVVLLLTFVSAALATQA.....DRKTYDIFNEECVYNGEPWL 209

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	282	25.1	171	13	US-10-085-572-7	Sequence 7, Appli
2	282	25.1	171	13	US-10-087-195-7	Sequence 7, Appli
3	251.5	22.4	182	13	US-10-085-572-8	Sequence 8, Appli
4	251.5	22.4	182	13	US-10-087-195-8	Sequence 8, Appli
5	239	21.2	172	13	US-10-085-572-6	Sequence 6, Appli
6	239	21.2	172	13	US-10-087-195-6	Sequence 6, Appli
7	102.5	9.1	722	9	US-09-815-242-10796	Sequence 10796, A
8	97.5	8.7	221	9	US-09-728-914-4	Sequence 4, Appli
9	84	7.5	217	9	US-09-728-914-22	Sequence 22, Appl
10	83	7.4	2285	9	US-09-932-183A-2	Sequence 2, Appli
11	82	7.3	381	15	US-10-369-493-5686	Sequence 5686, Ap
12	81.5	7.2	627	16	US-10-389-566-2357	Sequence 2357, Ap
13	81	7.2	659	13	US-10-090-624-12	Sequence 12, Appl
14	80.5	7.2	738	14	US-10-291-583-83	Sequence 83, Appl
15	79.5	7.1	738	14	US-10-291-583-92	Sequence 92, Appl
16	79.5	7.1	738	14	US-10-291-583-95	Sequence 95, Appl
17	79.5	7.1	738	15	US-10-423-704A-2	Sequence 2, Appli
18	79.5	7.1	913	14	US-10-043-344-5	Sequence 5, Appli
19	78.5	7.0	733	14	US-10-291-583-87	Sequence 87, Appl
20	78.5	7.0	738	14	US-10-291-583-80	Sequence 80, Appl
21	77.5	6.9	506	15	US-10-369-493-5064	Sequence 5064, Ap
22	77.5	6.9	685	14	US-10-291-583-113	Sequence 113, App
23	77.5	6.9	733	14	US-10-291-583-86	Sequence 86, Appl
24	77.5	6.9	733	14	US-10-291-583-90	Sequence 90, Appl
25	77.5	6.9	738	14	US-10-291-583-79	Sequence 79, Appl
26	77.5	6.9	738	14	US-10-291-583-81	Sequence 81, Appl
27	77.5	6.9	738	14	US-10-291-583-84	Sequence 84, Appl
28	77.5	6.9	738	14	US-10-291-583-85	Sequence 85, Appl
29	77.5	6.9	738	14	US-10-291-583-91	Sequence 91, Appl
30	77.5	6.9	738	14	US-10-291-583-93	Sequence 93, Appl
31	77.5	6.9	738	14	US-10-291-583-94	Sequence 94, Appl
32	77	6.8	631	14	US-10-032-585-7764	Sequence 7764, Ap
33	77	6.8	914	14	US-10-043-344-11	Sequence 11, Appl
34	76.5	6.8	413	16	US-10-380-077-18	Sequence 18, Appl
35	76.5	6.8	575	15	US-10-104-047-3622	Sequence 3622, Ap
36	76.5	6.8	738	14	US-10-291-583-82	Sequence 82, Appl
37	76.5	6.8	1778	14	US-10-238-075-749	Sequence 749, App
38	76	6.8	403	15	US-10-369-493-11803	Sequence 11803, A
39	75	6.7	518	10	US-09-769-734-56	Sequence 56, Appl
40	75	6.7	736	14	US-10-291-583-99	Sequence 99, Appl
41	75	6.7	912	14	US-10-043-344-7	Sequence 7, Appli
42	75	6.7	912	14	US-10-043-344-9	Sequence 9, Appli
43	74.5	6.6	381	15	US-10-369-493-21237	Sequence 21237, A
44	74.5	6.6	389	14	US-10-094-944-22	Sequence 22, Appl
45	74.5	6.6	492	15	US-10-369-493-23526	Sequence 23526, A

ALIGNMENTS

RESULT 1
US-10-085-572-7
; Sequence 7, Application US/10085572
; Publication No. US20020151499A1
; GENERAL INFORMATION:
; APPLICANT: Nuttall, Patricia, Ann
; APPLICANT: Paesen, Guido, Christiaan
; TITLE OF INVENTION: Treatment of Conjunctivitis
; FILE REFERENCE: 2488-1-003
; CURRENT APPLICATION NUMBER: US/10/085,572
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: PCT/GB00/03282
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 9920674.0
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 171
; TYPE: PRT
; ORGANISM: FS-HBP 2
US-10-085-572-7

Query Match 25.1%; Score 282; DB 13; Length 171;
Best Local Similarity 37.5%; Pred. No. 1.4e-22;
Matches 66; Conservative 29; Mismatches 69; Indels 12; Gaps 8;

```
Qy          31 PLWAHEELLGKYQDAWKSIDQGVSVTYVLAKTTYENDTGSWGSQFKCLQVQEIERKEEDY 90
               | || |   | :||||||:   |    | : | ||:||   ||: | |: |   :  |::
Db           3 PDWADEAANGAHQDAWKSLKADVENVYYMVKATYKNDP-VWGNDFTCVGVMANDVNEDEK 61

Qy          91 TVTSVFTF-RNASSPIKYYNVTETVKAVFQYGYKNIRNAIEYQVGGGLNITDTLIFTDGE 149
               :: : | |  || : :::   || | ||  ||| |  ||  |:   |   || : ::| :
Db          62 SIQAEFLFMNNADTNMQF--ATEKVTAVKMYGY-NRENAFRYETEDGQVFTDVIAYSD-D 117

Qy         150 LCDVFYVPNAD---QGCELWVKKSHYKHVPDYCTFVFNVFCAKDRKTYDIFNEECV 202
               ||| |||  |   :| |||   : | ::|  |   || : |  |:| |:|  |:
Db         118 NCDVIYVPGTDGNEEGYELWT--TDYDNIPANCLNKFNEY-AVGRETRDVFTSACL 170
```

RESULT 2
US-10-087-195-7
; Sequence 7, Application US/10087195
; Publication No. US20020193306A1
; GENERAL INFORMATION:
; APPLICANT: Nuttall, Patricia, Anne
; APPLICANT: Paesen, Guido, Christiaan
; TITLE OF INVENTION: Treatment of Allergic Rhinitis
; FILE REFERENCE: 2488-1-004
; CURRENT APPLICATION NUMBER: US/10/087,195
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: PCT/GB00/03287
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 9920673.2
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 171
; TYPE: PRT
; ORGANISM: FS-HBP 2
US-10-087-195-7

Query Match 25.1%; Score 282; DB 13; Length 171;
Best Local Similarity 37.5%; Pred. No. 1.4e-22;
Matches 66; Conservative 29; Mismatches 69; Indels 12; Gaps 8;

```
Qy          31 PLWAHEELLGKYQDAWKSIDQGVSVTYVLAKTTYENDTGSWGSQFKCLQVQEIERKEEDY 90
               | || |   | :|||||:  |  | : | ||:||   ||: | |: |   :  |::
Db           3 PDWADEAANGAHQDAWKSLKADVENVYYMVKATYKNDP-VWGNDFTCVGVMANDVNEDEK 61

Qy          91 TVTSVFTF-RNASSPIKYYNVTETVKAVFQYGYKNIRNAIEYQVGGGLNITDTLIFTDGE 149
               :: : | |  || : :::   || | ||  ||| |  ||  |:   |   || : ::| :
Db          62 SIQAEFLFMNNADTNMQF--ATEKVTAVKMYGY-NRENAFRYETEDGQVFTDVIAYSD-D 117

Qy         150 LCDVFYVPNAD---QGCELWVKKSHYKHVPDYCTFVFNVFCAKDRKTYDIFNEECV 202
               ||| |||  |   :| |||   : | ::|  |   || : |  :| |:|   |:
Db         118 NCDVIYVPGTDGNEEGYELWT--TDYDNIPANCLNKFNEY-AVGRETRDVFTSACL 170
```

RESULT 3
US-10-085-572-8
; Sequence 8, Application US/10085572
; Publication No. US20020151499A1
; GENERAL INFORMATION:
; APPLICANT: Nuttall, Patricia, Ann
; APPLICANT: Paesen, Guido, Christiaan
; TITLE OF INVENTION: Treatment of Conjunctivitis
; FILE REFERENCE: 2488-1-003
; CURRENT APPLICATION NUMBER: US/10/085,572
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: PCT/GB00/03282
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 9920674.0
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 182
; TYPE: PRT
; ORGANISM: MS-HBP 1
US-10-085-572-8

Query Match 22.4%; Score 251.5; DB 13; Length 182;
Best Local Similarity 35.3%; Pred. No. 3.3e-19;
Matches 61; Conservative 29; Mismatches 66; Indels 17; Gaps 8;

```
Qy          30 NPLWAHEELLGKYQDAWKSIDQGVSVTYVLAKTTYENDTGSWGSQFKCLQV--QEIERKE 87
               || ||:|  || |||||||: |  :  | ||: |   | | || :| |: |  ::| :|:
Db           1 NPTWANEAKLGSYQDAWKSLQQDQNKRYYLAQATQTTD-GVWGEEFTCVSVTAEKIGKKK 59

Qy          88 EDYTVTSVFTFRNAS-SPIKYYNVTETVKAVFQYGYKNIRNAIEYQVGGGLNIT--DTLI 144
               : |:    ::|   : :|  : | ||   : |      | |:|:  |    |  | :
Db          60 LNATI----LYKNKHLTDLKESHETITVWKAYDY---TTENGIKYETQGTRTQTFEDVFV 112

Qy         145 FTDGELCDVFYVP---NADQG-CELWVKKSHYKHVPDYCTFVFNVFCAKDRKT 193
               |:| : ||| :||   :|:|  |||| :    :|| | |  |  :  ||
Db         113 FSDYKNCDVIFVPKERGSDEGDYELWVSEDKIDKIPDCCKFTMAYFAQQQEKT 165
```

RESULT 4
US-10-087-195-8
; Sequence 8, Application US/10087195
; Publication No. US20020193306A1
; GENERAL INFORMATION:
; APPLICANT: Nuttall, Patricia, Anne
; APPLICANT: Paesen, Guido, Christiaan
; TITLE OF INVENTION: Treatment of Allergic Rhinitis
; FILE REFERENCE: 2488-1-004
; CURRENT APPLICATION NUMBER: US/10/087,195
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: PCT/GB00/03287
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 9920673.2
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 182
; TYPE: PRT
; ORGANISM: MS-HBP 1
US-10-087-195-8

Query Match 22.4%; Score 251.5; DB 13; Length 182;
Best Local Similarity 35.3%; Pred. No. 3.3e-19;
Matches 61; Conservative 29; Mismatches 66; Indels 17; Gaps 8;

```
Qy          30 NPLWAHEELLGKYQDAWKSIDQGVSVTYVLAKTTYENDTGSWGSQFKCLQV--QEIERKE 87
               || ||:|  || |||||||: |  :  | ||: |   | | || :| |: |  ::| :|:
Db           1 NPTWANEAKLGSYQDAWKSLQQDQNKRYYLAQATQTTD-GVWGEEFTCVSVTAEKIGKKK 59

Qy          88 EDYTVTSVFTFRNAS-SPIKYYNVTETVKAVFQYGYKNIRNAIEYQVGGGLNIT--DTLI 144
               : |:    ::|   : :|  : | ||   : |      | |:|:  |    |  | :
Db          60 LNATI----LYKNKHLTDLKESHETITVWKAYDY---TTENGIKYETQGTRTQTFEDVFV 112

Qy         145 FTDGELCDVFYVP---NADQG-CELWVKKSHYKHVPDYCTFVFNVFCAKDRKT 193
               |:| : ||| :||   :|:|  |||| :    :|| | |  |  :  ||
Db         113 FSDYKNCDVIFVPKERGSDEGDYELWVSEDKIDKIPDCCKFTMAYFAQQQEKT 165
```

RESULT 5
US-10-085-572-6
; Sequence 6, Application US/10085572
; Publication No. US20020151499A1
; GENERAL INFORMATION:
; APPLICANT: Nuttall, Patricia, Ann
; APPLICANT: Paesen, Guido, Christiaan
; TITLE OF INVENTION: Treatment of Conjunctivitis
; FILE REFERENCE: 2488-1-003
; CURRENT APPLICATION NUMBER: US/10/085,572
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: PCT/GB00/03282
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 9920674.0
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 172
; TYPE: PRT
; ORGANISM: FS-HBP1
US-10-085-572-6

Query Match 21.2%; Score 239; DB 13; Length 172;

```
  Best Local Similarity   34.5%;  Pred. No. 7.2e-18;
  Matches   61;  Conservative   24;  Mismatches   82;  Indels   10;  Gaps    7;

Qy          29 ENPLWAHEELLGKYQDAWKSIDQGVSVTYVLAKTTYENDTGSWGSQFKCLQVQEIERKEE 88
               : |:|| |    |::||||| : : |   | | | ||:||   ||: | |:        |:
Db           1 DKPVWADEAANGEHQDAWKHLQKLVEENYDLIKATYKNDP-VWGNDFTCVGTAAQNLNED 59

Qy          89 DYTVTSVFTFRNASSPIKYYNVTETVKAVFQYGYKNIRNAIEYQVGGGLNITDTLIFTDG 148
               :  | : | | | :  : | :  |         ||| |  ||| ||   |  :|| | |:|
Db          60 EKNVEAWFMFMNNADTV-YQHTFEKATPDKMYGY-NKENAITYQTEDGQVLTDVLAFSD- 116

Qy         149 ELCDVFYVPNAD---QGCELWVKKSHYKHVPDYCTFVFNVFCAKDRKTYDIFNEECV 202
               : | | |    |    | |||   : |  ||  |   || : |       |::  :|:
Db         117 DNCYVIYALGPDGSGAGYELWA--TDYTDVPASCLEKFNEYAA-GLPVRDVYTSDCL 170


RESULT 6
US-10-087-195-6
; Sequence 6, Application US/10087195
; Publication No. US20020193306A1
; GENERAL INFORMATION:
;  APPLICANT: Nuttall, Patricia, Anne
;  APPLICANT:  Paesen, Guido, Christiaan
;  TITLE OF INVENTION: Treatment of Allergic Rhinitis
;  FILE REFERENCE: 2488-1-004
;  CURRENT APPLICATION NUMBER: US/10/087,195
;  CURRENT FILING DATE:  2002-03-01
;  PRIOR APPLICATION NUMBER: PCT/GB00/03287
;  PRIOR FILING DATE: 2000-08-24
;  PRIOR APPLICATION NUMBER: 9920673.2
;  PRIOR FILING DATE: 1999-09-01
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 6
;   LENGTH: 172
;   TYPE: PRT
;   ORGANISM: FS-HBP1
US-10-087-195-6

  Query Match             21.2%;  Score 239;  DB 13;  Length 172;
  Best Local Similarity   34.5%;  Pred. No. 7.2e-18;
  Matches   61;  Conservative   24;  Mismatches   82;  Indels   10;  Gaps    7;

Qy          29 ENPLWAHEELLGKYQDAWKSIDQGVSVTYVLAKTTYENDTGSWGSQFKCLQVQEIERKEE 88
               : |:|| |    |::||||| : : |   | | | ||:||   ||: | |:        |:
Db           1 DKPVWADEAANGEHQDAWKHLQKLVEENYDLIKATYKNDP-VWGNDFTCVGTAAQNLNED 59

Qy          89 DYTVTSVFTFRNASSPIKYYNVTETVKAVFQYGYKNIRNAIEYQVGGGLNITDTLIFTDG 148
               :  | : | | | :  : | :  |         ||| |  ||| ||   |  :|| | |:|
Db          60 EKNVEAWFMFMNNADTV-YQHTFEKATPDKMYGY-NKENAITYQTEDGQVLTDVLAFSD- 116

Qy         149 ELCDVFYVPNAD---QGCELWVKKSHYKHVPDYCTFVFNVFCAKDRKTYDIFNEECV 202
               : | | |    |    | |||   : |  ||  |   || : |       |::  :|:
Db         117 DNCYVIYALGPDGSGAGYELWA--TDYTDVPASCLEKFNEYAA-GLPVRDVYTSDCL 170


RESULT 7
US-09-815-242-10796
; Sequence 10796, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 10796
;   LENGTH: 722
;   TYPE: PRT
;   ORGANISM: Enterococcus faecalis
US-09-815-242-10796

  Query Match             9.1%;  Score 102.5;  DB 9;  Length 722;
  Best Local Similarity   26.9%;  Pred. No. 0.042;
  Matches   60;  Conservative   26;  Mismatches   72;  Indels   65;  Gaps   15;

Qy          16 LATQAETTSA-----KAGENPLWAHEELLGKYQDAWKSI-DQGVSVTYVLAKTTY----- 64
               |   || |:      | |:||    ||: |:     |:: | |    ||      |
Db         204 LVGDKETGSSYGFAVKKGQNP-----ELIKKFNAGLKNLKDNG---TYDKILNNYLATGD 255

Qy          65 ENDTGSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASSPIKYYNV--------TETVKA 116
               | :|   | | |     :|  |:| | : |  ||    :| :: |          : ||
Db         256 ETNTQDAGEQMK-----KITPKKEKYVIASDSTF----APFEFQNAQGDYVGIDVDLVKR 306

Qy         117 V---------FQY-GYKNIRNAIEY-QVGG---GLNITDTLIFTDGELCDVFYVPNADQG 162
                       |:: |: :   |:|  |  |   |: |||      | :    | ||  | |
Db         307 AAELQGFTVEFKFIGFSSAVQAVESGQADGMVAGMTITD-----DRKKAFDFSVPYFDSG 361

Qy         163 CELWVKKSH--YKHVPDYCTFVFNV--------FCAKDRKTYD 195
                :: ||| :    |    |         |    |::| ||
Db         362 IQIAVKKGNDKIKSYDDLKGKKVGVKIGTESADFLEKNKKKYD 404


RESULT 8
US-09-728-914-4
; Sequence 4, Application US/09728914
; Patent No. US20010046499A1
; GENERAL INFORMATION:
;  APPLICANT: KANTOR, FRED S.
;  APPLICANT:  FIKRIG, EROL
;  APPLICANT:  DAS, SUBRATA
;  TITLE OF INVENTION: TICK ANTIGENS AND COMPOSITIONS AND METHODS COMPRISING
;  TITLE OF INVENTION:  THEM
;  FILE REFERENCE: YU-107
;  CURRENT APPLICATION NUMBER: US/09/728,914
;  CURRENT FILING DATE:  2000-12-01
;  PRIOR APPLICATION NUMBER: 60/169,048
;  PRIOR FILING DATE: 1999-12-03
;  PRIOR APPLICATION NUMBER: 60/240,716
;  PRIOR FILING DATE: 2000-10-16
;  NUMBER OF SEQ ID NOS: 39
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 221
;   TYPE: PRT
;   ORGANISM: Ixodes scapularis
US-09-728-914-4

  Query Match             8.7%;  Score 97.5;  DB 9;  Length 221;
  Best Local Similarity   23.4%;  Pred. No. 0.028;
  Matches   49;  Conservative   43;  Mismatches   80;  Indels   37;  Gaps   14;
```

```
Qy          10 TFVSAALATQAET-------TSAKAGENPLWAHEELLGKY--QDAWKSIDQGVSVTYVLA 60
               |: |    |:  |       |:|   |:|          ||  |:| : ::  :: |  :
Db          31 TYGSTGTTTRPGTRGARMIVTTAPPEEDP--------SKYKEQNATRVVE--MNATQWVK 80

Qy          61 KTTYENDTGSWGSQFKCLQVQEIERKEEDYTVTSV-FTFRNASSPIKYYNVTET--VKAV 117
                ||:  |   |:  :|   : :|::    | |:  | :|  |    :  : ||  :|: :
Db          81 WRTYD-VTDFSGNPVQCENFRVMEKR----TPTNYSFQYRYKSKN-SWETIDETLILKDI 134

Qy         118 FQYGYKNIRNAIEYQVGGGLNITDTLI-FTDGELCDVFYVPNADQG---CELWVKK-SHY 172
               ::|: |: |     :|     || |: :::   | |  :|  :||   |:||:   :
Db         135 GEHGFPNVMNFQRTPIG---IATDNLVLYSNYVNCTVLRIPFTNQGERHCDLWMANLTLS 191

Qy         173 KHVPDYCTFVFNVFCAKDRKTYDIFNEEC 201
               :  || |   |  :|    : | ::   |
Db         192 QETPDDCLNKFFEYC-NTTQIYRVYYPSC 219
```

RESULT 9
US-09-728-914-22
; Sequence 22, Application US/09728914
; Patent No. US20010046499A1
; GENERAL INFORMATION:
; APPLICANT: KANTOR, FRED S.
; APPLICANT: FIKRIG, EROL
; APPLICANT: DAS, SUBRATA
; TITLE OF INVENTION: TICK ANTIGENS AND COMPOSITIONS AND METHODS COMPRISING
; TITLE OF INVENTION: THEM
; FILE REFERENCE: YU-107
; CURRENT APPLICATION NUMBER: US/09/728,914
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/169,048
; PRIOR FILING DATE: 1999-12-03
; PRIOR APPLICATION NUMBER: 60/240,716
; PRIOR FILING DATE: 2000-10-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Ixodes scapularis
US-09-728-914-22

Query Match 7.5%; Score 84; DB 9; Length 217;
Best Local Similarity 24.1%; Pred. No. 0.82;
Matches 27; Conservative 17; Mismatches 48; Indels 20; Gaps 3;

```
Qy          99 RNASSPIKYYNVTETVKAVFQYGYKNIRNAIEYQVGGGLNITDTLIFTDG---------- 148
               || ||    |   :   :    | :  |: || :: |       :  | |
Db          79 RNDSS----YEFVKRTLSWISVGEQPSRHTIEVEITGTSKSKSEVKVTKGYEEYDFNTVY 134

Qy         149 --ELCDVFYVPNAD----QGCELWVKKSHYKHVPDYCTFVFNVFCAKDRKTY 194
                 | :  :       : | |||||:  |:   :| |:|:|||   |: :
Db         135 ADTRCLILRISRTQKVPLRSCLLWVKKTFLKNPLRHCRFLFDVFCNWRREDF 186
```

RESULT 10
US-09-932-183A-2
; Sequence 2, Application US/09932183A
; Patent No. US20020127641A1
; GENERAL INFORMATION:
; APPLICANT: Estell, David A.
; TITLE OF INVENTION: Proteases From Gram-Positive Organisms
; FILE REFERENCE: GC394C1-US
; CURRENT APPLICATION NUMBER: US/09/932,183A
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 09/308,375
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: PCT/US98/18828
; PRIOR FILING DATE: 1998-09-08
; PRIOR APPLICATION NUMBER: EP9719636.4
; PRIOR FILING DATE: 1997-09-15
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 2285
; TYPE: PRT
; ORGANISM: Bacillius subtilis
US-09-932-183A-2

Query Match 7.4%; Score 83; DB 9; Length 2285;
Best Local Similarity 23.3%; Pred. No. 28;
Matches 51; Conservative 27; Mismatches 79; Indels 62; Gaps 10;

```
Qy           6 VLLLTFVSAALATQAETTSAKAGENPLWAHEELLGKYQDAWKSID--QGVSVTYVLAKTT 63
               :|: | |  | |              || | |:  : :| |: |  :   | | | ||
Db         743 LLVSTLVGGAFAALG-----------WALESLISSFAEAKKAKDDFEQSQQTNVEAITT 790

Qy          64 YENDTGSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASSPIKYYNVTETVKAVFQ---Y 120
                :: |     |:|     |::: :|  ::||          :|  ||: :   |
Db         791 NKDSTDKLIQQYK-----ELQKVKESRSLTS-------DEEQEYLQVTQQLAQTFPALVK 838

Qy         121 GY-----------KNIRNAIEYQVGGGLNITDTLIFTDGELCDVFYVPNADQGCELWVKK 169
               ||           | :  |||       |  : |     |  |       |   | :||
Db         839 GYDSQGNAILKTNKELEKAIE-------NTKEYLALKKQETRDSAKKTFEDASKE--IKK 889

Qy         170 S-----HYKHVPDYCTFVFNVFCAKDRKTYDIFNEECVY 203
               |     || : ||          | |  :|:  ::  |
Db         890 SKDELKQYKQIADYND--------KGRPKWDLIADDDDY 920
```

RESULT 11
US-10-369-493-5686
; Sequence 5686, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5686
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-5686

Query Match 7.3%; Score 82; DB 15; Length 381;
Best Local Similarity 21.8%; Pred. No. 3;
Matches 29; Conservative 20; Mismatches 56; Indels 28; Gaps 3;

```
Qy           9 LTFVSAALATQAETTSAKAGENPLWAHEELLG----------KYQDAWKSIDQGVSVTY 57
               |  |    |  || |: ::  | | | | :|            |: | |    : |:  :
Db          35 LEIVDQTKAASAEETNNESEVNELDADEPMLQDLDNRFVIFPLKHHDIWNFYKKAVASFW 94

Qy          58 VLAKTTYENDTGSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASSPI-------KYYNV 110
                : :     |   |           |    :| | :: :  |  ||  |        :: |
Db          95 TVEEVDLGKDMNDW----------EKMNGDEQYFISRILAFFAASDGIVNENLCERFSNE 144

Qy         111 TETVKAVFQYGYK 123
                :  :| | ||::
Db         145 VQVSEARFFYGFQ 157
```

RESULT 12
US-10-389-566-2357
; Sequence 2357, Application US/10389566
; Publication No. US20040025202A1
; GENERAL INFORMATION:

```
;  APPLICANT: Monsanto Technology, LLC
;  APPLICANT:  Laurie, Cathy C
;  TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
;  FILE REFERENCE: 38-77(52900)D
;  CURRENT APPLICATION NUMBER: US/10/389,566
;  CURRENT FILING DATE:  2003-03-31
;  PRIOR APPLICATION NUMBER: US 60/365,301
;  PRIOR FILING DATE: 2002-03-15
;  PRIOR APPLICATION NUMBER: US 60/391,786
;  PRIOR FILING DATE: 2002-06-25
;  PRIOR APPLICATION NUMBER: US 60/392,018
;  PRIOR FILING DATE: 2002-06-26
;  NUMBER OF SEQ ID NOS: 2459
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2357
;   LENGTH: 627
;   TYPE: PRT
;   ORGANISM: Rickettsia prowazekii
US-10-389-566-2357

  Query Match             7.2%;  Score 81.5;  DB 16;  Length 627;
  Best Local Similarity   28.2%;  Pred. No. 6.7;
  Matches   33;  Conservative   15;  Mismatches   42;  Indels   27;  Gaps    7;

Qy         72 GSQFKCLQVQE------IERKEEDYTVTSVFTFRNA----SSPIKYYNVTETVKAVFQYG 121
              |:   |: | |       |:  | : |  |:  | |:       | |    ||
Db         10 GTTNSCVAVMEGKEPKVIDNAEGERTTPSIIAFANSERLVGQPAKRQAVT---------- 59

Qy        122 YKNIRNAIEYQVGG--GLNITDTLIFTDGELCDVFYVPNADQGCELWVKKSHYKHVP 176
               | || | | |    | | || ::  | |   : :  || | : ||:  ::|: |
Db         60 --NPRNTI-YAVKRLIGRNFTDPMVRKDQGLVP-YNIVKADNG-DAWVEADNHKYSP 111


RESULT 13
US-10-090-624-12
; Sequence 12, Application US/10090624
; Publication No. US20020132335A1
; GENERAL INFORMATION:
;  APPLICANT: TAKAKURA, Hikaru
;  APPLICANT:  MORISHITA, Mio
;  APPLICANT:  SHIMOJO, Tomoko
;  APPLICANT:  ASADA, Kiyozo
;  APPLICANT:  KATO, Ikunoshin
;  TITLE OF INVENTION: SYSTEM FOR EXPRESSING HYPERTHERMOSTABLE
;  FILE REFERENCE: TAKAKURA=6
;  CURRENT APPLICATION NUMBER: US/10/090,624
;  CURRENT FILING DATE:  2002-03-06
;  PRIOR APPLICATION NUMBER: 09/445,472
;  PRIOR FILING DATE: 1999-12-06
;  PRIOR APPLICATION NUMBER: 151969/1997
;  PRIOR FILING DATE: 1997-06-10
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
;   LENGTH: 659
;   TYPE: PRT
;   ORGANISM: Thermococcus celer
US-10-090-624-12

  Query Match             7.2%;  Score 81;  DB 13;  Length 659;
  Best Local Similarity   25.9%;  Pred. No. 8.2;
  Matches   53;  Conservative   20;  Mismatches   62;  Indels   70;  Gaps   13;

Qy         41 KYQDAWK------SIDQGVSV-TYVLAKTTYENDTGSW--GSQFKCLQVQEIERKEEDYT 91
              || |  |       |:| :  |: ::  |:   |  | ||    | : :    | ||:
Db        436 KYDDYAKLTFTGSVADKGSATHTFDVSGATFVTATLYWDTGSSDIDLYLYDPNGNEVDYS 495

Qy         92 VTSVFTFRNASSPIKYYNVTE---TVKAVFQYGYKNIRNAIEYQV----------GGG- 136
              |: : |      : ||| |    ||| |    ||   |   |||           |||
Db        496 YTAYYGFEK----VGYYNPTAGTWTVKVV---SYKGAAN---YQVDVVSDGSLSQSGGGN 545

Qy        137 ---------LNITDTLIFTDGELCDVFYVPNADQGCELWVKKSHYKHVPDYCTFVFNVFC 187
                       |||  || | : |  |   :|                  ||  ||

Db        546 PNPNPNPNPTPTTDTQTFT-GSVND--YWDTSD-----------------TFTMNVNS 583

Qy        188 AKDRKTYDI-----FNEECVYNGEP 207
                : | |:      :|:  :|  :|
Db        584 GATKITGDLTFDTSYNDLDLYLYDP 608


RESULT 14
US-10-291-583-83
; Sequence 83, Application US/10291583
; Publication No. US20030138772A1
; GENERAL INFORMATION:
;  APPLICANT: Gao, Guangping
;  APPLICANT:  Wilson, James M.
;  APPLICANT:  Alvira, Mauricio
;  TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Vir
;  TITLE OF INVENTION:  Sequences and Isolating No. US20030138772A1el Sequences Iden
;  FILE REFERENCE: UPN-02735USA
;  CURRENT APPLICATION NUMBER: US/10/291,583
;  CURRENT FILING DATE:  2002-11-12
;  PRIOR APPLICATION NUMBER: US 60/350,607
;  PRIOR FILING DATE: 2001-11-13
;  PRIOR APPLICATION NUMBER: US 60/341,117
;  PRIOR FILING DATE: 2001-12-17
;  PRIOR APPLICATION NUMBER: US 60/377,066
;  PRIOR FILING DATE: 2002-05-01
;  PRIOR APPLICATION NUMBER: US 60/386,675
;  PRIOR FILING DATE: 2002-06-05
;  NUMBER OF SEQ ID NOS: 120
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 83
;   LENGTH: 738
;   TYPE: PRT
;   ORGANISM: capsid protein of AAV serotype, clone 29.5VP1
US-10-291-583-83

  Query Match             7.2%;  Score 80.5;  DB 14;  Length 738;
  Best Local Similarity   27.5%;  Pred. No. 11;
  Matches   41;  Conservative   19;  Mismatches   56;  Indels   33;  Gaps   10;

Qy         22 TTSAKAGENPLWAHEELLGKYQD-AWKSIDQGVSVTYVLAKTTYENDTGSWG----SQFK 76
              ||| :      ||  |  | :  :| |  | |          ||   :  ||     ::|
Db        242 TTSTRT-----WA----LPTYNNHLYKQISNGTSGGST-NDNTYFGYSTPWGYFDFNRFH 291

Qy         77 CLQVQEIERKEEDYTVTSVFTFRNASSPIKYYNVTETVKAVFQ-YGYKNIRNAIEYQVGG 135
              |       ::    : : : ||  |    | :|:   || | |  | | | |
Db        292 C----HFSPRDWQRLINNNWGFRPKSLNFKLFNI--QVKEVTQNEGTKTIAN-------- 337

Qy        136 GLNITDTL-IFTDGELCDVFYVPNADQGC 163
                |:| |: :||| |    : : :| |||
Db        338 --NLTSTIQVFTDSEYQLPYVLGSAHQGC 364


RESULT 15
US-10-291-583-92
; Sequence 92, Application US/10291583
; Publication No. US20030138772A1
; GENERAL INFORMATION:
;  APPLICANT: Gao, Guangping
;  APPLICANT:  Wilson, James M.
;  APPLICANT:  Alvira, Mauricio
;  TITLE OF INVENTION: A Method of Detecting and/or Identifying Adeno-Associated Vir
;  TITLE OF INVENTION:  Sequences and Isolating No. US20030138772A1el Sequences Iden
;  FILE REFERENCE: UPN-02735USA
;  CURRENT APPLICATION NUMBER: US/10/291,583
;  CURRENT FILING DATE:  2002-11-12
;  PRIOR APPLICATION NUMBER: US 60/350,607
;  PRIOR FILING DATE: 2001-11-13
;  PRIOR APPLICATION NUMBER: US 60/341,117
;  PRIOR FILING DATE: 2001-12-17
;  PRIOR APPLICATION NUMBER: US 60/377,066
;  PRIOR FILING DATE: 2002-05-01
;  PRIOR APPLICATION NUMBER: US 60/386,675
```

; PRIOR FILING DATE: 2002-06-05
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 92
; LENGTH: 738
; TYPE: PRT
; ORGANISM: capsid protein of AAV serotype, clone 43.1
US-10-291-583-92

Query Match 7.1%; Score 79.5; DB 14; Length 738;
Best Local Similarity 28.3%; Pred. No. 14;
Matches 45; Conservative 19; Mismatches 60; Indels 35; Gaps 12;

```
Qy          13 SAALATQAETTSAKAGENPLWAHEELLGKYQD-AWKSIDQGVSVTYVLAKTTYENDTGSW 71
               |  |  :  ||| :      ||     |  | :  :| |  | |         ||   :  |
Db         233 STWLGDRVITTSTRT-----WA----LPTYNNHLYKQISNGTSGGST-NDNTYFGYSTPW 282

Qy          72 G----SQFKCLQVQEIERKEEDYTVTSVFTFRNASSPIKYYNVTETVKAVFQ-YGYKNIR 126
               |    ::| |        ::    : : : ||      | :|:   || | |  | | |
Db         283 GYFDFNRFHC----HFSPRDWQRLINNNWGFRPKRLNFKLFNI--QVKEVTQNEGTKTIA 336

Qy         127 NAIEYQVGGGLNITDTL-IFTDGELCDVFYVP-NADQGC 163
               |          |:| |: :||| |   : ||| :| |||
Db         337 N----------NLTSTIQVFTDSEY-QLPYVPGSAHQGC 364
```

Search completed: March 9, 2004, 13:36:50
Job time : 34 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 9, 2004, 13:30:31 ; Search time 20 Seconds
(without alignments)
1005.201 Million cell updates/sec

Title: US-09-555-296B-4
Perfect score: 1125
Sequence: 1 MKMQVVLLLTFVSAALATQA.....DRKTYDIFNEECVYNGEPWL 209

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	96.5	8.6	621	2	F97862	heat shock protein
2	96	8.5	650	2	G70169	heat shock protein
3	87	7.7	639	2	F71258	probable heat shoc
4	86	7.6	2747	2	B49132	fat facets (faf) s
5	85	7.6	656	2	A87112	heat shock protein
6	84.5	7.5	621	2	H71645	heat shock protein
7	84.5	7.5	917	2	AF1394	autolysin, amidase
8	83	7.4	2285	2	T12796	probable transglyc
9	82.5	7.3	318	2	C69192	stomatin-like prot
10	82	7.3	163	2	C97308	probable acetyltra
11	82	7.3	360	2	T17996	hypothetical prote
12	82	7.3	381	2	T18876	hypothetical prote
13	82	7.3	390	2	D82922	conserved hypothet
14	82	7.3	637	2	T08530	traG protein - Ent
15	82	7.3	637	2	S22992	traG protein - Esc
16	81.5	7.2	627	2	D71729	dnaK-type molecula
17	80.5	7.2	621	2	A71961	90Kda chaperone -
18	79.5	7.1	621	2	B64546	chaperone and heat
19	79.5	7.1	725	2	AG2547	hypothetical prote
20	78.5	7.0	455	2	A75121	hypothetical prote
21	78.5	7.0	864	2	S60441	hypothetical prote
22	78	6.9	215	2	A26669	nodulin-22 precurs
23	78	6.9	523	1	O4CKA3	cytochrome P450 52
24	78	6.9	1871	2	S27938	hypothetical prote
25	77.5	6.9	223	2	T12076	ribonuclease (EC 3
26	77.5	6.9	506	2	T25870	hypothetical prote
27	77	6.8	241	2	T26909	hypothetical prote
28	77	6.8	331	2	D81409	probable periplasm
29	77	6.8	780	2	H84453	probable heat shoc
30	77	6.8	914	2	S70906	transferrin-bindin
31	76.5	6.8	587	2	E87420	poly-beta-hydroxyb
32	76.5	6.8	635	2	T46407	probable RNA helic
33	76.5	6.8	807	2	T18454	hypothetical prote
34	76.5	6.8	2819	2	A90551	conserved hypothet
35	76	6.8	260	2	A55122	deoxycytidine kina
36	76	6.8	510	2	A71816	hypothetical prote
37	76	6.8	510	2	A64706	lipase-like protei
38	76	6.8	531	2	T18741	hypothetical prote
39	76	6.8	598	2	AG0890	probable arylsulfa
40	76	6.8	781	2	S49340	heat-shock protein
41	76	6.8	1031	2	T38411	probable GTPase ac
42	76	6.8	1066	2	B95037	hyaluronidase [imp
43	76	6.8	1068	2	S73091	hypothetical prote
44	76	6.8	1277	2	B84517	hypothetical prote
45	75.5	6.7	820	2	D81382	probable transmemb

ALIGNMENTS

RESULT 1
F97862
heat shock protein htpG [imported] - Rickettsia conorii (strain Malish 7)
C;Species: Rickettsia conorii
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 22-Oct-2001
C;Accession: F97862
R;Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.
Science 293, 2093-2098, 2001
A;Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A;Reference number: A97700; MUID:21442074; PMID:11557893
A;Accession: F97862
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-621 <KUR>
A;Cross-references: GB:AE006914; PIDN:AAL03840.1; PID:g15620441; GSPDB:GN00173
C;Genetics:
A;Gene: htpG
C;Superfamily: heat shock protein 90

Query Match 8.6%; Score 96.5; DB 2; Length 621;
Best Local Similarity 21.0%; Pred. No. 1.2;
Matches 43; Conservative 46; Mismatches 81; Indels 35; Gaps 10;

```
Qy          11 FVSAALATQAETTSAKAGENP--LWAHEELLGKY--QDAWKSIDQGVSVTYVLAKTTYEN 66
               :|  :| :   || ||||:   :|   : ||:|   |: |   :|  :   :|   |
Db         127 YSSFMVADKVTVTSRKAGEDKVHIW-ESDGLGEYTVSDSDKEFTRGTEIVLHIKK---EE 182

Qy          67 DTGSWGSQFKCLQVQEIERKEEDYTVTSVFTF---------RNASSPIKYYNVTETVKAV 117
               ||  :   |:   :: | :   |:    :: |          |::| :    :|  :
Db         183 DT--FLDHFR---LKHIVKSYSDHIAVPIYFFDEAGNNEIQLNSASALWTRPKSEITEEQ 237

Qy         118 FQYGYKNIRNAIE-------YQVGGGLNITDTLIFTDGELCDVFYVPNADQGCELWVKK- 169
               ::  ||:: ||:         :  | :  |: |      :  |:|: |:  :  :|::|:
Db         238 YKEFYKSLSYAIDDPWITMHNKNEGAIEFTNLLFIPSSKTFDLFH-PDRKRRVKLYIKRV 296

Qy         170 ----SHYKHVPDYCTFVPNVFCAKD 190
                    :| |  |:  |  ::|
Db         297 FISDENIDLIPSYLRFLRGVVDSED 321
```

RESULT 2
G70169
heat shock protein 90 (htpG) homolog - Lyme disease spirochete
N;Alternate names: C62.5 heat shock protein
C;Species: Borrelia burgdorferi (Lyme disease spirochete)
C;Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #text_change 20-Aug-1999
C;Accession: G70169; I40247
R;Fraser, C.M.; Casjens, S.; Huang, W.M.; Sutton, G.G.; Clayton, R.; Lathigra, R.; W
son, D.; Peterson, J.; Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.; V
; Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst, K.; Roberts, K.; Hatch, B
Nature 390, 580-586, 1997
A;Authors: Smith, H.O.; Venter, J.C.

A;Title: Genomic sequence of a Lyme disease spirochaete, Borrelia burgdorferi.
A;Reference number: A70100; MUID:98065943; PMID:9403685
A;Accession: G70169
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-650 <KLE>
A;Cross-references: GB:AE001157; GB:AE000783; NID:g2688471; PIDN:AAC66919.1; PID:g268847
A;Experimental source: strain B31
R;Ojaimi, C.; Davidson, B.E.; Saint Girons, I.; Old, I.G.
Microbiology 140, 2931-2940, 1994
A;Title: Conservation of gene arrangement and an unusual organization of rRNA genes in t
A;Reference number: I40241; MUID:95111614; PMID:7812434
A;Accession: I40247
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 531-609 <RES>
A;Cross-references: GB:L32145; NID:g476004; PIDN:AAC41403.1; PID:g476594
C;Genetics:
A;Gene: htpG
C;Superfamily: heat shock protein 90

Query Match 8.5%; Score 96; DB 2; Length 650;
Best Local Similarity 21.5%; Pred. No. 1.4;
Matches 46; Conservative 34; Mismatches 92; Indels 42; Gaps 8;

```
Qy          11 FVSAALATQ-AETTSAKAGENP--LWAHEELLGKYQDAWKSIDQGVSVTYVLAKTTYEND 67
               | || : ::  | || || |:  :|: :   |   :  |  : |  :   | |   |
Db         155 FYSAFIVSEKVEVTSKKALESDAYIWSSDGKTGYEIEKAKKEESGTEIKLYLNKEGLE-- 212

Qy          68 TGSWGSQFKCLQV-------------------------QEIERKEEDYTVTSVFTFRNA 101
                  : :::|  ::                        : || |||    |:    :|
Db         213 ---YANKWKIQEIIKKYSNHINYPIYIKYSEPIMKDGKQEGIEEKEEKLNETTALWTKN- 268

Qy         102 SSPIKYYNVTETVKAVFQYGYKNIRNAIEYQVGGGLNITDTLIFTDGELCDVFYVPNADQ 161
                | ||     |  |    : |:|     |  :  | |  |:          |::| ||
Db         269 KSEIKAEEYNEFYKNT-TFDYENPLMHIHTKAEGNLEYTNLFYVPSKAPYDLYY-PNTKP 326

Qy         162 GCELWVKK-----SHYKHVPDYCTFVFNVFCAKD 190
               | :|:: :     |    :|:|  |:  :   :|
Db         327 GVKLFINRIFITDSEGSLLPNYLRFIKGIIDCQD 360
```

RESULT 3
F71258
probable heat shock protein 90 (htpG) - syphilis spirochete
C;Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Nov-1999
C;Accession: F71258
R;Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDo
they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A;Reference number: A71250; MUID:98332770; PMID:9665876
A;Accession: F71258
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-639 <COL>
A;Cross-references: GB:AE001265; GB:AE000520; NID:g3323295; PIDN:AAC65938.1; PID:g332330
A;Experimental source: strain Nichols
C;Genetics:
A;Gene: TP0984
C;Superfamily: heat shock protein 90

Query Match 7.7%; Score 87; DB 2; Length 639;
Best Local Similarity 22.9%; Pred. No. 9;
Matches 50; Conservative 26; Mismatches 52; Indels 90; Gaps 10;

```
Qy           5 VVLLLTFVSAALATQAETTSAKAGENPLWAHEELLGKYQD----------AWKSIDQGVS 54
               ||| |:  ::  ||:             |  ||:: || |          |  |:  :
Db         182 VVLHLSQENSEFATR-------------WRLEEVIKKYSDHIAFPIYLHYLQKEYDKDGA 228

Qy          55 VTYVLAKTTYENDTGSWGSQFKCLQVQEIERKEEDY-----TVTSVFTFRNASSPIKYYN 109
```

```
               ||    |    || |:   : |      | |||||     |:|     |::: |: | :
Db         229 VTDTQKKVDQVNDAGALWKRPKS------ELKEEDYHRFYQTLT-----RDSTPPLLYVH 277

Qy         110 VTETVKAVFQYGYKNIRNAIEYQVGGGLNITDTLIFTDG--ELCDVFYVP--------NA 159
                                                  :|  |   :||||        :|
Db         278 TK---------------------------------AEGTQEYVTLFYVPAKAPFDLFHA 303

Qy         160 D--QGCELWVKK-----SHYKHVPDYCTFVFNVFCAKD 190
               |   | :|:||:        : :| |  ||  |  ::|
Db         304 DYKPGVKLFVKRVFITDDEKELLPVYLRFVRGVIDSED 341
```

RESULT 4
B49132
fat facets (faf) splice form 1 - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 19-Dec-1993 #sequence_revision 25-Apr-1997 #text_change 01-Dec-2000
C;Accession: B49132; A49132
R;Fischer-Vize, J.A.; Rubin, G.M.; Lehmann, R.
Development 116, 985-1000, 1992
A;Title: The fat facets gene is required for Drosophila eye and embryo development.
A;Reference number: A49132; MUID:93202020; PMID:1295747
A;Contents: isogenic st
A;Accession: B49132
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-2747 <FIS>
A;Cross-references: GB:L04959; NID:g157411; PIDN:AAF01345.1; PID:g6013474
A;Note: sequence inconsistent with the nucleotide translation
A;Note: sequence extracted from NCBI backbone (NCBIN:127836, NCBIN:129008, NCBIP:12
A;Accession: A49132
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-2704,'VT',2707,'ANNV' <FI2>
A;Cross-references: GB:L04958; NID:g157410; PIDN:AAF01346.1; PID:g6013475
A;Note: sequence extracted from NCBI backbone (NCBIN:127836, NCBIN:129008, NCBIP:12
C;Keywords: alternative splicing

Query Match 7.6%; Score 86; DB 2; Length 2747;
Best Local Similarity 22.9%; Pred. No. 75;
Matches 54; Conservative 28; Mismatches 72; Indels 82; Gaps 12;

```
Qy           6 VLLLTFVSAALATQAETTSAKAGENPLWAHEELLG------KYQDA-------WKSIDQG 52
               |::|  | |  |     :          || | :|||      : |||        :|:|:|
Db        1756 VVILKHVQAIFAHLGHSALQYYVPRGLWTHFKLLGEPVNLREQQDAVEFFMSLLESLDEG 1815

Qy          53 VSVTYVLAKTTYENDT--GSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASSPIKYYNV 110
               :     | :    | |  ||:  |  | :      ||| ::| ||   || ||      :
Db        1816 LK---ALGQPQLMNATLGGSFSDQKICQECPHRYSKEEPFSVFSV-DIRNHSS------L 1865

Qy          111 TETVKAVFQYGYKNIRNAIEYQVGGGLNITDTLIFTDGEL---CDVFYVPNADQGCELWV 167
                ||:::                          :  |||    | ::    |:   : |
Db        1866 TESLEQ---------------------------YVKGELLEGADAYHCDKCDKKV-VTV 1896

Qy          168 KKSHYKHVP-----------DY---CTFVFNVFCAKDRKTYDIFNEECVYNGEPW 208
                |:   | :|            ||   |  ||          | |    : : ||:
Db        1897 KRVCVKKLPPVLAIQLKRFEYDYERVCAIKFN----------DYFEFPRILDMEPY 1942
```

RESULT 5
A87112
heat shock protein Hsp90 family [imported] - Mycobacterium leprae
C;Species: Mycobacterium leprae
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C;Accession: A87112
R;Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Hol
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R
A;Title: Massive gene decay in the leprosy bacillus.
A;Reference number: A86909; MUID:21128732; PMID:11234002
A;Accession: A87112

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-656 <STO>
A;Cross-references: GB:AL450380; NID:g13093407; PIDN:CAC30574.1; GSPDB:GN00147
C;Genetics:
A;Gene: htpG
C;Superfamily: heat shock protein 90

Query Match 7.6%; Score 85; DB 2; Length 656;
Best Local Similarity 21.3%; Pred. No. 14;
Matches 49; Conservative 38; Mismatches 79; Indels 64; Gaps 9;

```
Qy         11 FVSAALATQAETTSAKAGENPL--WAHEELLGKYQDAWKSID---QGVSVTYVLAKTTYE 65
Db        137 YSSFMVANKVELLTRKAGETAATRWSSD---GEATYTIESVDEAPQGTSVTLHLKPEDFE 193

Qy         66 NDTGSWGSQFKCLQV-------------QEIERK------------EEDYTVTSVFTFRN 100
Db        194 DELHDYTSEWKIRELVKKYSDFIAWPIRMEVERRAPATSDGEGADGEEQVTIETQTI--N 251

Qy        101 ASSPIKYYNVTETVKAVFQYGYKNIRNA--------------IEYQVGGGLNITDTLIF 145
Db        252 SMKALWTKSKDEVSEDEYKEFYKHIAHAWDDPLEVIAMKAEGTFEYQ---------ALLF 302

Qy        146 TDGELCDVFYVPNADQGCELWVKKSHY-----KHVPDYCTFVFNVFCAKD 190
Db        303 IPSHAPFDLFNSDAKIGMQLYVKRVFIMSDCDQLMPMYLRFVKGVVDAED 352
```

RESULT 6
H71645
heat shock protein htpG (htpG) RP840 - Rickettsia prowazekii
C;Species: Rickettsia prowazekii
C;Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 03-Nov-2000
C;Accession: H71645
R;Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark, U
Nature 396, 133-140, 1998
A;Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.
A;Reference number: A71630; MUID:99039499; PMID:9823893
A;Accession: H71645
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-621 <AND>
A;Cross-references: GB:AJ235273; GB:AJ235269; NID:g3861237; PIDN:CAA15264.1; PID:g386136
A;Experimental source: strain Madrid E
C;Genetics:
A;Gene: htpG; RP840
C;Superfamily: heat shock protein 90

Query Match 7.5%; Score 84.5; DB 2; Length 621;
Best Local Similarity 18.9%; Pred. No. 15;
Matches 39; Conservative 42; Mismatches 88; Indels 37; Gaps 8;

```
Qy         11 FVSAALATQAETTSAKAGENPLWAHE-ELLGKY--QDAWKSIDQGVSVTYVLAK--TTYE 65
Db        127 YSSFMVADKVTVTSRKAGESKVHTWESDGLGEYIVADSEQEFTRGTEIVLYIKKSETTFL 186

Qy         66 NDTGSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASSPIKYYNV-----TETVKAVFQY 120
Db        187 DH----------FRLKHIVKSYSDHIAVPIYFCDEAGNNEIQLNSASALWTRPKSEITED 236

Qy        121 GYKNIRNAIEYQVG-----------GGLNITDTLIFTDGELCDVFYVPNADQGCELWVKK 169
Db        237 QYKEFYKSLSYAVDDPWVTLHNKNEGAIEFTNLLFIPSSKTFDLFH-PDRKKRVKLYIKR 295

Qy        170 -----SHYKHVPDYCTFVFNVFCAKD 190
Db        296 VFISDENIDLIPSYLRFLRGVVDSED 321
```

RESULT 7
AF1394
autolysin, amidase [imported] - Listeria monocytogenes (strain EGD-e)
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C;Accession: AF1394
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloe
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehl
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AF1394
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-917 <GLA>
A;Cross-references: GB:NC_003210; PIDN:CAD00636.1; PID:g16412046; GSPDB:GN00177
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: ami

Query Match 7.5%; Score 84.5; DB 2; Length 917;
Best Local Similarity 22.6%; Pred. No. 24;
Matches 37; Conservative 25; Mismatches 49; Indels 53; Gaps 9;

```
Qy         28 GENPLWAHEELLGKYQDAWKSIDQGVSVTYVLAKTTYENDT---GSWG----------S 73
Db        223 GKGTVWSHD----------------AVTRYLGGTTHTDPVAYFNQWGYNFNNFVSLINE 265

Qy         74 QFKCLQVQEIERKEEDYTVTSVFTFRNASSPIKYY--NVTETVKAV---FQYGYKNIR-- 126
Db        266 KYKAMQV-NYEKIEYDKAITAYSRVKTATGNSVWTKPNKTEGAKLVNPLSSYSGKNLRII 324

Qy        127 -------NAIEYQ--VGGGLNITDTLIFTDGELCDVFYVPNADQ 161
Db        325 REAKTSGGTIWYQFSIGG-----KTIGWVDSKALNTFYTPSMEK 363
```

RESULT 8
T12796
probable transglycosylase - Bacillus subtilis phage SPBc2
C;Species: Bacillus subtilis phage SPBc2
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 15-Oct-1999
C;Accession: T12796; A69911
R;Lazarevic, V.; Duesterhoeft, A.; Soldo, B.; Hilbert, H.; Mauel, C.; Karamata, D.
submitted to the EMBL Data Library, August 1997
A;Description: The complete nucleotide sequence of the Bacillus subtilis SPbetac2 pr
A;Reference number: Z17583
A;Accession: T12796
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-2285 <LAZ>
A;Cross-references: EMBL:AF020713; NID:g3025478; PID:g3025510; PIDN:AAC13005.1
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Be
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Ga
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo,
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardir
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Ma
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portet
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scar
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; S
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiy
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshic
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtil
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: A69911
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-2285 <KUN>
A;Cross-references: GB:Z99115; GB:AL009126; NID:g2634478; PIDN:CAB14053.1; PID:e1183

A;Experimental source: strain 168
C;Genetics:
A;Gene: yomI

Query Match 7.4%; Score 83; DB 2; Length 2285;
Best Local Similarity 23.3%; Pred. No. 1.1e+02;
Matches 51; Conservative 27; Mismatches 79; Indels 62; Gaps 10;

```
Qy          6 VLLLTFVSAALATQAETTSAKAGENPLWAHEELLGKYQDAWKSID--QGVSVTYVLAKTT 63
Db        743 LLVSTLVGGAFAALG-----------WALESLISSFAEAKKAKDDFEQSQQTNVEAITT 790

Qy         64 YENDTGSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASSPIKYYNVTETVKAVFQ---Y 120
Db        791 NKDSTDKLIQQYK-----ELQKVKESRSLTS-------DEEQEYLQVTQQLAQTFPALVK 838

Qy        121 GY-----------KNIRNAIEYQVGGGLNITDTLIFTDGELCDVFYVPNADQGCELWVKK 169
Db        839 GYDSQGNAILKTNKELEKAIE-------NTKEYLALKKQETRDSAKKTFEDASKE--IKK 889

Qy        170 S-----HYKHVPDYCTFVFNVFCAKDRKTYDIFNEECVY 203
Db        890 SKDELKQYKQIADYND--------KGRPKWDLIADDDDY 920
```

RESULT 9
C69192
stomatin-like protein - Methanobacterium thermoautotrophicum (strain Delta H)
C;Species: Methanobacterium thermoautotrophicum
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 24-Sep-1999
C;Accession: C69192
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
ki, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A;Reference number: A69000; MUID:98037514; PMID:9371463
A;Accession: C69192
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-318 <MTH>
A;Cross-references: GB:AE000848; GB:AE000666; NID:g2621761; PIDN:AAB85197.1; PID:g262177
A;Experimental source: strain Delta H
C;Genetics:
A;Gene: MTH692
C;Superfamily: erythrocyte band 7 integral membrane protein

Query Match 7.3%; Score 82.5; DB 2; Length 318;
Best Local Similarity 23.3%; Pred. No. 9.2;
Matches 38; Conservative 23; Mismatches 77; Indels 25; Gaps 6;

```
Qy          5 VVLLLTFVSAALATQAETTSAKAGENPLWAHEELLGKYQDAWKSIDQGVSVTYVLAKTTY 64
Db         10 VLLVLAFKSLKILRPYEKGVV----------ERLGKYQ---RTVESGLVVIIPFIEAIK 55

Qy         65 ENDTGSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASSPIK-YYNVTETVKAVFQYGYK 123
Db         56 KVD---MREQVVDVPPQEVITKDNTVVVVDCVIFYEVVDPFNAVYNVVDFYQAITKLAQT 112

Qy        124 NIRNAIEYQVGGGLNITDTLIFTDGELCDVFYVPNADQGCELW 166
Db        113 NLRNII-----GDLELDQTL--TSREMINTQLREVLDEATDKW 148
```

RESULT 10
C97308
probable acetyltransferase [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C;Accession: C97308
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacteriu
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: C97308
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-163 <KUR>
A;Cross-references: GB:AE001437; PIDN:AAK81254.1; PID:g15026402; GSPDB:GN00168
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC3322

Query Match 7.3%; Score 82; DB 2; Length 163;
Best Local Similarity 24.5%; Pred. No. 4.2;
Matches 37; Conservative 11; Mismatches 51; Indels 52; Gaps 6;

```
Qy         82 EIERKEEDYTVTSVFTFRNASSPIKYYNVTETVKAV----------------FQYGYKNI 125
Db          3 KISVKQMDYETAKQITKWNYEKPFSIYNMNESDECIKELLSGYYFSAYDDKNYIVGYYGF 62

Qy        126 RNAIEYQVGGGL------NITDTLIFTDGELCDVFYVPNADQG-------CELWVKKSHY 172
Db         63 GEAAQIPVGEQLGIYSYKNITDIGIGLNPNLC--------GQGLGFDFFCCGL------- 107

Qy        173 KHVPDYCTFVFNVFCAKD-RKTYDIFNEECV 202
Db        108 -------DFARNTLCAKDFRLTVATFNKPAI 131
```

RESULT 11
T17996
hypothetical protein A494R - Chlorella virus PBCV-1
C;Species: Chlorella virus PBCV-1
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T17996
R;Graves, M.V.; Van Etten, J.L.
submitted to the EMBL Data Library, May 1999
A;Reference number: Z18806
A;Accession: T17996
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-360 <GRA>
A;Cross-references: EMBL:U42580; NID:g4028896; PIDN:AAC96861.1
A;Experimental source: specific host Chlorella strain NC64A
C;Genetics:
A;Note: A494R

Query Match 7.3%; Score 82; DB 2; Length 360;
Best Local Similarity 25.8%; Pred. No. 12;
Matches 41; Conservative 19; Mismatches 59; Indels 40; Gaps 8;

```
Qy         35 HEELLGK---YQDAWKSIDQGVSVTYVLAKTTYENDTGSWGSQFKCLQVQEIERKEEDYT 91
Db         43 HEAVNGNIQLFQEKKMKLENGVN-----TKKTEE------------KIKQIEEDEINYL 84

Qy         92 VTSVFTFRNASSP-IKYYNVTETV------KAVFQYGYKNIR-NAIE---YQVGGGLNIT 140
Db         85 LDTI--------PFIKEYDVKETVSDVTEQNSVFQVKSKNTHTNTFRKYLFHVEKVSNPT 136

Qy        141 DTLIFTDGELCDVFYVPNADQGCELWVKKSHYKHVPDYC 179
Db        137 TLDAVTDRETVDQIYTCTCGGQMELWVNSTQSDLVCNEC 175
```

RESULT 12
T18876
hypothetical protein C03C10.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Sep-2003
C;Accession: T18876
R;Berks, M.
submitted to the EMBL Data Library, August 1994
A;Reference number: Z19036
A;Accession: T18876

A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-381 <WIL>
A;Cross-references: EMBL:Z35637; PIDN:CAA84688.1; GSPDB:GN00021; CESP:C03C10.3
A;Experimental source: clone C03C10
C;Genetics:
A;Gene: CESP:C03C10.3
A;Map position: 3
A;Introns: 18/3; 137/3; 256/2
C;Superfamily: ribonucleoside reductase small subunit

Query Match 7.3%; Score 82; DB 2; Length 381;
Best Local Similarity 21.8%; Pred. No. 13;
Matches 29; Conservative 20; Mismatches 56; Indels 28; Gaps 3;

```
Qy          9 LTFVSAALATQAETTSAKAGENPLWAHEELLG-----------KYQDAWKSIDQGVSVTY 57
              |  |      |  || |: ::  | | | | :|            |: | |    : |:  :
Db         35 LEIVDQTKAASAEETNNESEVNELDADEPMLQDLDNRFVIFPLKHHDIWNFYKKAVASFW 94

Qy         58 VLAKTTYENDTGSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASSPI-------KYYNV 110
               : :      |   |           |    :| | :: :  |  ||  |       :: |
Db         95 TVEEVDLGKDMNDW----------EKMNGDEQYFISRILAFFAASDGIVNENLCERFSNE 144

Qy        111 TETVKAVFQYGYK 123
              :  :| | ||::
Db        145 VQVSEARFFYGFQ 157
```

RESULT 13
D82922
conserved hypothetical UU198 [imported] - Ureaplasma urealyticum
C;Species: Ureaplasma urealyticum
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C;Accession: D82922
R;Glass, J.I.; Lefkowitz, E.J.; Glass, J.S.; Heiner, C.R.; Chen, E.Y.; Cassell, G.H.
submitted to GenBank, February 2000
A;Description: The complete sequence of Ureaplasma urealyticum: Alternate views of a min
A;Reference number: A82870
A;Accession: D82922
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-390 <GLA>
A;Cross-references: GB:AE002119; GB:AF222894; NID:g6899153; PIDN:AAF30605.1; GSPDB:GN001
A;Experimental source: serovar 3; biovar 1
C;Genetics:
A;Gene: UU198
A;Genetic code: SGC3

Query Match 7.3%; Score 82; DB 2; Length 390;
Best Local Similarity 22.3%; Pred. No. 13;
Matches 29; Conservative 28; Mismatches 51; Indels 22; Gaps 5;

```
Qy          8 LLTFVSAALATQAETTSAKAGENPLWAHEELLGKYQDAWKSIDQGVSVTY-VLAKTTYEN 66
              :| :      | : | |     ||| : :|   || | :| |: |   ||  | ::| |:
Db        167 ILAYSRTGGANASGTAEAFLTSNPLISKKEFKKKYHDIYKVIEAG---TYGKLIRSTEES 223

Qy         67 DTGSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASSPIKYYNVTETVKAVFQYGY---K 123
              :| :|           | :   | ::| :       |:      |  : :   ::  |   |
Db        224 NTQAW---------TNISKDNVDGSITYL------STGFILKNYDQIIAKGYKIAYLQNK 268

Qy        124 NIRNAIEYQV 133
              :::| : :::
Db        269 DLKNYLPFEI 278
```

RESULT 14
T08530
traG protein - Enterobacter aerogenes plasmid R751
C;Species: Enterobacter aerogenes
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C;Accession: T08530
R;Thorsted, P.B.; Shah, D.S.; Macartney, D.; Kostelidou, K.; Thomas, C.M.
Plasmid 36, 95-111, 1996
A;Title: Conservation of the genetic switch between replication and transfer genes c
A;Reference number: Z16434; MUID:97118926; PMID:8954881
A;Accession: T08530
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-637 <THO>
A;Cross-references: EMBL:U67194; NID:g1572520; PIDN:AAC64474.1; PID:g1572579
C;Genetics:
A;Gene: traG
A;Genome: plasmid R751

Query Match 7.3%; Score 82; DB 2; Length 637;
Best Local Similarity 18.3%; Pred. No. 25;
Matches 46; Conservative 30; Mismatches 89; Indels 86; Gaps 7;

```
Qy          5 VVLLLTFVSAALATQAETTSAKAGE----NPLWAHEE---------------------- 37
              :|  :  :  |:|    : |:|| |    :  || ::
Db         90 LVSTVGLLGVAVAKVVTSNSSKANEYLHGSARWAEKKDIQAAGLLPRERNVLEIVTGKAA 149

Qy         38 --LLGKYQDAWKSID--------QGVSVTYVLAKTTYENDTG-------SWGSQFKCLQV 80
                 | |   |:  |         |      | |       |        |||:     :
Db        150 PTATGVYVGGWQDKDGNFFYLRHSGPEHVLTYAPTRSGKGVGLVVPTLLSWGASSVITDL 209

Qy         81 QEIERKEEDYTVTSVFTFRNASSPIKYYNVTETVKAVFQYGYKNIRNAIEYQVGGGLNIT 140
                 | | : :|: :  ::| : :  :    |   |          ||   ||:||    |:
Db        210 -----KGELWALTAGWRQKHAKNKVLRFEPASTSGGVCWNPLDEIRLGTEYEVGDVQNLA 264

Qy        141 DTLIPTDGELCDVFYV-----------------------------------PNADQGCE 164
                ::  ||:  |  :                                     || | | |
Db        265 TLIVDPDGKGLDSHWQKTAFALLVGVILHALYKAKDDGGTATLPSVDAMLADPNRDIG-E 323

Qy        165 LWVKKSHYKHV 175
              ||:: : | ||
Db        324 LWMEMATYGHV 334
```

RESULT 15
S22992
traG protein - Escherichia coli plasmid R751
C;Species: Escherichia coli
C;Date: 31-Dec-1993 #sequence_revision 02-Aug-1994 #text_change 08-Oct-1999
C;Accession: S22992; S42274
R;Ziegelin, G.; Pansegrau, W.; Strack, B.; Balzer, D.; Kroeger, M.; Kruft, V.; Lanka
DNA Seq. 1, 303-327, 1991
A;Title: Nucleotide sequence and organization of genes flanking the transfer origin
A;Reference number: S22992; MUID:92190548; PMID:1665997
A;Accession: S22992
A;Molecule type: DNA
A;Residues: 1-637 <ZIE>
A;Cross-references: EMBL:X54458
R;Lanka, E.
submitted to the EMBL Data Library, August 1990
A;Reference number: S42274
A;Accession: S42274
A;Molecule type: DNA
A;Residues: 1-628,'D',630-637 <LAN>
A;Cross-references: EMBL:X54458; NID:g42656; PIDN:CAA38327.1; PID:g42657
C;Genetics:
A;Gene: traG
A;Genome: plasmid

Query Match 7.3%; Score 82; DB 2; Length 637;
Best Local Similarity 18.3%; Pred. No. 25;
Matches 46; Conservative 30; Mismatches 89; Indels 86; Gaps 7;

```
Qy          5 VVLLLTFVSAALATQAETTSAKAGE----NPLWAHEE---------------------- 37
              :|  :  :  |:|    : |:|| |    :  || ::
Db         90 LVSTVGLLGVAVAKVVTSNSSKANEYLHGSARWAEKKDIQAAGLLPRERNVLEIVTGKAA 149

Qy         38 --LLGKYQDAWKSID--------QGVSVTYVLAKTTYENDTG-------SWGSQFKCLQV 80
                 | |   |:  |         |      | |       |        |||:     :
Db        150 PTATGVYVGGWQDKDGNFFYLRHSGPEHVLTYAPTRSGKGVGLVVPTLLSWGASSVITDL 209
```

```
Qy          81 QEIERKEEDYTVTSVFTFRNASSPIKYYNVTETVKAVFQYGYKNIRNAIEYQVGGGLNIT 140
               | | : :|: :  ::| : :  :    |   |       ||   ||:||   |:
Db         210 -----KGELWALTAGWRQKHAKNKVLRFEPASTSGGVCWNPLDEIRLGTEYEVGDVQNLA 264

Qy         141 DTLIFTDGELCDVFYV----------------------------------PNADQGCE 164
                 ::  ||:  |  :                                  || | | |
Db         265 TLIVDPDGKGLDSHWQKTAFALLVGVILHALYKAKDDGGTATLPSVDAMLADPNRDIG-E 323

Qy         165 LWVKKSHYKHV 175
               ||:: : | ||
Db         324 LWMEMATYGHV 334
```

Search completed: March 9, 2004, 13:34:01
Job time : 22 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 9, 2004, 13:30:56 ; Search time 17 Seconds
(without alignments)
640.157 Million cell updates/sec

Title: US-09-555-296B-4
Perfect score: 1125
Sequence: 1 MKMQVVLLLTFVSAALATQA.....DRKTYDIFNEECVYNGEPWL 209

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	290.5	25.8	190	1	HBP2_RHIAP	O77421 rhipicephal
2	265	23.6	200	1	HBPM_RHIAP	O77422 rhipicephal
3	252	22.4	190	1	HBP1_RHIAP	O77420 rhipicephal
4	97.5	8.7	174	1	NUGC_MESVI	Q9mur1 mesostigma
5	96.5	8.6	621	1	HTPG_RICCN	P58478 rickettsia
6	96	8.5	616	1	HTPG_BORBU	P42555 borrelia bu
7	87	7.7	639	1	HTPG_TREPA	O83949 treponema p
8	86	7.6	2778	1	FAF_DROME	P55824 drosophila
9	85	7.6	656	1	HTPG_MYCLE	O33012 mycobacteri
10	84.5	7.5	621	1	HTPG_RICPR	Q9zcb9 rickettsia
11	82.5	7.3	318	1	Y692_METTH	O26788 methanobact
12	82	7.3	381	1	RIR2_CAEEL	P42170 caenorhabdi
13	82	7.3	637	1	TRG5_ECOLI	Q00184 escherichia
14	81.5	7.2	627	1	DNAK_RICPR	Q9zdx9 rickettsia
15	81	7.2	700	1	HS9C_DICDI	P54651 dictyosteli
16	80.5	7.2	621	1	HTPG_HELPJ	Q9zmm2 helicobacte
17	79.5	7.1	621	1	HTPG_HELPY	P56116 helicobacte
18	79.5	7.1	629	1	HTPG_CHLTE	Q8ke61 chlorobium
19	79.5	7.1	725	1	FTKL_ANASP	Q8zs46 anabaena sp
20	78.5	7.0	864	1	YG3M_YEAST	P48237 saccharomyc
21	78	6.9	215	1	NO22_SOYBN	P08961 glycine max
22	77	6.8	523	1	CP5C_CANMA	P16496 candida mal
23	77	6.8	909	1	HEX_ADEM1	P48308 mouse adeno
24	76.5	6.8	393	1	IDH_STRMU	Q59940 streptococc
25	76.5	6.8	724	1	DDX4_HUMAN	Q9nqi0 homo sapien
26	76	6.8	260	1	DCK_MOUSE	P43346 mus musculu
27	76	6.8	1031	1	YDG9_SCHPO	Q10496 schizosacch
28	76	6.8	1066	1	HYSA_STRPN	Q54873 streptococc
29	76	6.8	1068	1	TRI_SULSO	P95871 sulfolobus
30	75.5	6.7	1489	1	RNG2_SCHPO	O14188 schizosacch
31	75	6.7	414	1	Y878_METJA	Q58288 methanococc
32	75	6.7	424	1	PORA_CAMJE	P80672 campylobact
33	75	6.7	566	1	BCDO_RAT	Q91xt5 rattus norv
34	75	6.7	626	1	HTPG_BUCBP	Q89a93 buchnera ap
35	75	6.7	647	1	HTPG_MYCTU	Q50667 mycobacteri
36	75	6.7	6629	1	R1AB_IBVB	P27920 a replicase
37	75	6.7	6629	1	R1AB_IBVBC	Q91qt2 a replicase
38	74.5	6.6	400	1	GTR3_RABIT	Q9xsc2 oryctolagus
39	74.5	6.6	492	1	YGCE_ECOLI	P55138 escherichia
40	74.5	6.6	726	1	NFI1_YEAST	Q12216 saccharomyc
41	74.5	6.6	890	1	LPN1_HUMAN	Q14693 homo sapien
42	74.5	6.6	2768	1	THYG_HUMAN	P01266 homo sapien
43	74	6.6	406	1	IDH_SPHYA	P50215 sphingomona
44	74	6.6	782	1	POT6_ARATH	Q8w4i4 arabidopsis
45	74	6.6	854	1	TRIC_SULTO	Q972g5 sulfolobus

ALIGNMENTS

RESULT 1
HBP2_RHIAP
ID HBP2_RHIAP STANDARD; PRT; 190 AA.
AC O77421;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Female-specific histamine-binding protein 2 precursor (FS-HBP2).
OS Rhipicephalus appendiculatus (Brown ear tick).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Parasitiformes; Ixodida; Ixodidae; Rhipicephalus.
OX NCBI_TaxID=34631;
RN [1]
RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (1.25 ANGSTROMS).
RC TISSUE=Salivary gland;
RX MEDLINE=99288454; PubMed=10360182;
RA Paesen G.C., Adams P.L., Harlos K., Nuttall P.A., Stuart D.I.;
RT "Tick histamine-binding proteins: isolation, cloning, and three-dimensional structure.";
RL Mol. Cell 3:661-671(1999).
CC -!- FUNCTION: Binds histamine with a high-affinity. The ability to outcompete histamine receptors indicates that its function is to suppress inflammation during blood feeding.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC ---

CC ---
DR EMBL; U96081; AAC63107.1; -.
DR PDB; 1QFT; 19-APR-00.
DR PDB; 1QFV; 19-APR-00.
DR InterPro; IPR002970; His_binding.
DR Pfam; PF02098; His_binding; 1.
DR ProDom; PD152455; His_binding; 1.
KW Signal; 3D-structure.
FT SIGNAL 1 19
FT CHAIN 20 190 FEMALE-SPECIFIC HISTAMINE-BINDING PROTEIN 2.
FT DISULFID 67 188
FT DISULFID 138 167
FT TURN 23 24
FT HELIX 27 30
FT HELIX 31 33
FT HELIX 36 41
FT TURN 42 45
FT STRAND 48 53
FT STRAND 57 58
FT TURN 59 61
FT STRAND 62 63
FT TURN 64 64
FT STRAND 66 76

```
FT   TURN         77     80
FT   STRAND       81     89
FT   TURN         91     92
FT   STRAND       97    106
FT   TURN        109    110
FT   STRAND      116    121
FT   TURN        122    123
FT   STRAND      126    135
FT   TURN        136    137
FT   STRAND      138    143
FT   STRAND      152    157
FT   TURN        160    161
FT   HELIX       165    174
FT   TURN        175    177
FT   STRAND      181    182
FT   TURN        186    188
SQ   SEQUENCE   190 AA;  21464 MW;  6923A3E902552B6F CRC64;

  Query Match             25.8%;  Score 290.5;  DB 1;  Length 190;
  Best Local Similarity   36.0%;  Pred. No. 1e-18;
  Matches   72;  Conservative   32;  Mismatches   79;  Indels   17;  Gaps    9;

Qy          7 LLLTFVSAALATQAETTSAKAGENPLWAHEELLGKYQDAWKSIDQGVSVTYVLAKTTYEN 66
              ||:  ::  ||        |   |  | || |   | :||||||:   |   | : | ||:|
Db          3 LLILSLALVLA-----LSQVKGNQPDWADEAANGAHQDAWKSLKADVENVYYMVKATYKN 57

Qy         67 DTGSWGSQFKCLQVQEIERKEEDYTVTSVFTF-RNASSPIKYYNVTETVKAVFQYGYKNI 125
              |   ||: | |: |   :  |:: :: : | |  || : :::   || | ||  ||| |
Db         58 DP-VWGNDFTCVGVMANDVNEDEKSIQAEFLFMNNADTNMQF--ATEKVTAVKMYGY-NR 113

Qy        126 RNAIEYQVGGGLNITDTLIFTDGELCDVFYVPNAD---QGCELWVKKSHYKHVPDYCTFV 182
               ||  |:    |   || : ::| : ||| |||  |   :| |||   : | ::|  |
Db        114 ENAFRYETEDGQVFTDVIAYSD-DNCDVIYVPGTDGNEEGYELWT--TDYDNIPANCLNK 170

Qy        183 FNVFCAKDRKTYDIFNEECV 202
              || : |  |:| |:|   |:
Db        171 FNEY-AVGRETRDVFTSACL 189


RESULT 2
HBPM_RHIAP
ID   HBPM_RHIAP     STANDARD;      PRT;   200 AA.
AC   O77422;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   Male-specific histamine-binding salivary protein precursor (MS-HBP).
OS   Rhipicephalus appendiculatus (Brown ear tick).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC   Parasitiformes; Ixodida; Ixodidae; Rhipicephalus.
OX   NCBI_TaxID=34631;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Salivary gland;
RX   MEDLINE=99288454; PubMed=10360182;
RA   Paesen G.C., Adams P.L., Harlos K., Nuttall P.A., Stuart D.I.;
RT   "Tick histamine-binding proteins: isolation, cloning, and three-
RT   dimensional structure.";
RL   Mol. Cell 3:661-671(1999).
CC   -!- FUNCTION: Binds histamine with a high-affinity. The ability to
CC       outcompete histamine receptors indicates that its function is to
CC       suppress inflammation during blood feeding.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U96082; AAC63108.1; -.
DR   HSSP; O77421; 1QFT.
DR   InterPro; IPR002970; His_binding.
DR   Pfam; PF02098; His_binding; 1.
DR   ProDom; PD152455; His_binding; 1.
KW   Signal.
FT   SIGNAL       1     18       POTENTIAL.
FT   CHAIN       19    200       MALE-SPECIFIC HISTAMINE-BINDING SALIVARY
FT                               PROTEIN.
FT   DISULFID     65    193       BY SIMILARITY.
FT   DISULFID    137    169       BY SIMILARITY.
SQ   SEQUENCE   200 AA;  22851 MW;  C46A1C8C6BCAA008 CRC64;

  Query Match             23.6%;  Score 265;  DB 1;  Length 200;
  Best Local Similarity   34.5%;  Pred. No. 2e-16;
  Matches   69;  Conservative   34;  Mismatches   71;  Indels   26;  Gaps   10;

Qy          3 MQVVLLLTFVSAALATQAETTSAKAGENPLWAHEELLGKYQDAWKSIDQGVSVTYVLAKT 62
              |:|:||:  : |||   |:          || ||:|  || |||||||: |  :  | ||:
Db          1 MKVLLLV--LGAALCQNADA-------NPTWANEAKLGSYQDAWKSLQQDQNKRYYLAQA 51

Qy         63 TYENDTGSWGSQFKCLQV--QEIERKEEDYTVTSVFTFRNAS-SPIKYYNVTETVKAVFQ 119
              |   | | || :| |: |  ::| :|: : |:     ::|   : :|  : | ||   :
Db         52 TQTTD-GVWGEEFTCVSVTAEKIGKKKLNATI----LYKNKHLTDLKESHETITVWKAYD 106

Qy        120 YGYKNIRNAIEYQVGGGLNIT--DTLIFTDGELCDVFYVP---NADQG-CELWVKKSHYK 173
              |      | |:|:  |   |  |  :|:| : ||| :||    :|:|  |||| :
Db        107 Y---TTENGIKYETQGTRTQTFEDVFVFSDYKNCDVIFVPKERGSDEGDYELWVSEDKID 163

Qy        174 HVPDYCTFVFNVFCAKDRKT 193
              :|| | |   |  :  ||
Db        164 KIPDCCKFTMAYFAQQQEKT 183


RESULT 3
HBP1_RHIAP
ID   HBP1_RHIAP     STANDARD;      PRT;   190 AA.
AC   O77420;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Female-specific histamine-binding protein 1 precursor (FS-HBP1).
OS   Rhipicephalus appendiculatus (Brown ear tick).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC   Parasitiformes; Ixodida; Ixodidae; Rhipicephalus.
OX   NCBI_TaxID=34631;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Salivary gland;
RX   MEDLINE=99288454; PubMed=10360182;
RA   Paesen G.C., Adams P.L., Harlos K., Nuttall P.A., Stuart D.I.;
RT   "Tick histamine-binding proteins: isolation, cloning, and three-
RT   dimensional structure.";
RL   Mol. Cell 3:661-671(1999).
CC   -!- FUNCTION: Binds histamine with a high-affinity. The ability to
CC       outcompete histamine receptors indicates that its function is to
CC       suppress inflammation during blood feeding.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U96080; AAC63106.1; -.
DR   HSSP; O77421; 1QFT.
DR   InterPro; IPR002970; His_binding.
DR   Pfam; PF02098; His_binding; 1.
DR   ProDom; PD152455; His_binding; 1.
KW   Signal.
FT   SIGNAL       1     18       POTENTIAL.
```

```
FT   CHAIN        19    190       FEMALE-SPECIFIC HISTAMINE-BINDING PROTEIN
FT                                1.
FT   DISULFID     66    187       BY SIMILARITY.
FT   DISULFID    137    166       BY SIMILARITY.
SQ   SEQUENCE   190 AA;  21370 MW;  855BE151A90053B1 CRC64;

  Query Match             22.4%;  Score 252;  DB 1;  Length 190;
  Best Local Similarity   33.5%;  Pred. No. 2.6e-15;
  Matches   68;  Conservative   28;  Mismatches   89;  Indels   18;  Gaps    8;

Qy          3 MQVVLLLTFVSAALATQAETTSAKAGENPLWAHEELLGKYQDAWKSIDQGVSVTYVLAKT 62
              |:::| | || |    :|        : |:|| |   |::||||| : : |   | | |
Db          1 MKLLLSLAFVLALSQVKA--------DKPVWADEAANGEHQDAWKHLQKLVEENYDLIKA 52

Qy         63 TYENDTGSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASSPIKYYNVTETVKAVFQYGY 122
              ||:||   ||: | |:        |::  | : | | | :  : | :  |       |||
Db         53 TYKNDP-VWGNDFTCVGTAAQNLNEDEKNVEAWFMFMNNADTV-YQHTFEKATPDKMYGY 110

Qy        123 KNIRNAIEYQVGGGLNITDTLIFTDGELCDVFYVPNAD---QGCELWVKKSHYKHVPDYC 179
              |  ||| ||   |  :|| | |:| : | | |     |   | |||   : |  ||  |
Db        111 -NKENAITYQTEDGQVLTDVLAFSD-DNCYVIYALGPDGSGAGYELWA--TDYTDVPASC 166

Qy        180 TFVFNVFCAKDRKTYDIFNEECV 202
              || : |      |::  :|:
Db        167 LEKFNEYAA-GLPVRDVYTSDCL 188


RESULT 4
NUGC_MESVI
ID   NUGC_MESVI     STANDARD;      PRT;   174 AA.
AC   Q9MUR1;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   NAD(P)H-quinone oxidoreductase chain J, chloroplast (EC 1.6.5.-)
DE   (NAD(P)H dehydrogenase, chain J) (NADH-plastoquinone oxidoreductase
DE   subunit J).
GN   NDHJ.
OS   Mesostigma viride.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Mesostigmatophyceae;
OC   Mesostigmatales; Mesostigmataceae; Mesostigma.
OX   NCBI_TaxID=41882;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NIES-296;
RX   MEDLINE=20150907; PubMed=10688199;
RA   Lemieux C., Otis C., Turmel M.;
RT   "Ancestral chloroplast genome in Mesostigma viride reveals an early
RT   branch of green plant evolution.";
RL   Nature 403:649-652(2000).
CC   -!- CATALYTIC ACTIVITY: NAD(P)H + plastoquinone = NAD(P)(+) +
CC       plastoquinol.
CC   -!- SIMILARITY: Belongs to the complex I 30 kDa subunit family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF166114; AAF43839.1; -.
DR   InterPro; IPR001268; Complex1_30K.
DR   Pfam; PF00329; complex1_30Kd; 1.
DR   ProDom; PD001581; Complex1_30K; 1.
DR   PROSITE; PS00542; COMPLEX1_30K; 1.
KW   Oxidoreductase; NAD; NADP; Quinone; Plastoquinone; Chloroplast.
SQ   SEQUENCE   174 AA;  20329 MW;  122ABFA9188D47B7 CRC64;

  Query Match             8.7%;  Score 97.5;  DB 1;  Length 174;
  Best Local Similarity  27.0%;  Pred. No. 0.11;
  Matches   31;  Conservative   20;  Mismatches   37;  Indels   27;  Gaps    5;

Qy        105 IKYYNVTETVKAVFQYGYKNIRNAIEYQVGGGLNITDTLIFTDGELCDVFYV----PNAD 160
              :|  |:|     |:: ||: :|:  | |  |              |:|  |:::    |||
Db         47 VKAQNLTSVAIALYAYGFNYLRSQCAYDVSPG-----------GDLASVYHLTKVDDNAD 95

Qy        161 QGCELWVKKSHYKHVPDYCTFVFNVFCA------KDRKTYDIFNEECVYNGEPWL 209
              |  |: :|      ||     : :||        ::|::||:|     | | | |
Db         96 QPQEVCIK----VFVPRTKPIIPSVFWIWKTADFQERESYDMFG--IYYEGHPHL 144


RESULT 5
HTPG_RICCN
ID   HTPG_RICCN     STANDARD;      PRT;   621 AA.
AC   P58478;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Chaperone protein htpG (Heat shock protein htpG) (High temperature
DE   protein G).
GN   HTPG OR RC1302.
OS   Rickettsia conorii.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=781;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Malish 7;
RX   MEDLINE=21442074; PubMed=11557893;
RA   Ogata H., Audic S., Renesto-Audiffren P., Fournier P.-E., Barbe V.,
RA   Samson D., Roux V., Cossart P., Weissenbach J., Claverie J.-M.,
RA   Raoult D.;
RT   "Mechanisms of evolution in Rickettsia conorii and R. prowazekii.";
RL   Science 293:2093-2098(2001).
CC   -!- FUNCTION: Molecular chaperone. Has ATPase activity (By
CC       similarity).
CC   -!- SUBUNIT: Homodimer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Belongs to the heat shock protein 90 family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE008676; AAL03840.1; -.
DR   PIR; F97862; F97862.
DR   HAMAP; MF_00505; -; 1.
DR   InterPro; IPR003594; ATPbind_ATPase.
DR   InterPro; IPR001404; Hsp90.
DR   Pfam; PF02518; HATPase_c; 1.
DR   Pfam; PF00183; HSP90; 2.
DR   PRINTS; PR00775; HEATSHOCK90.
DR   SMART; SM00387; HATPase_c; 1.
DR   PROSITE; PS00298; HSP90; 1.
KW   Chaperone; ATP-binding; Heat shock; Complete proteome.
FT   DOMAIN       1    328       A; SUBSTRATE-BINDING (BY SIMILARITY).
FT   DOMAIN     329    544       B (BY SIMILARITY).
FT   DOMAIN     545    621       C.
SQ   SEQUENCE   621 AA;  70804 MW;  FF6E5B352DD8F017 CRC64;

  Query Match             8.6%;  Score 96.5;  DB 1;  Length 621;
  Best Local Similarity  21.0%;  Pred. No. 0.63;
  Matches   43;  Conservative   46;  Mismatches   81;  Indels   35;  Gaps   10;

Qy         11 FVSAALATQAETTSAKAGENP--LWAHEELLGKY--QDAWKSIDQGVSVTYVLAKTTYEN 66
              : |  :| :   || ||||:   :|   :  ||:|   |: |   :|  :   : |  |
Db        127 YSSFMVADKVTVTSRKAGEDKVHIW-ESDGLGEYTVSDSDKEFTRGTEIVLHIKK---EE 182

Qy         67 DTGSWGSQFKCLQVQEIERKEEDYTVTSVFTF---------RNASSPIKYYNVTETVKAV 117
```

```
                ||  :   |:   :: | :   |:    :: |          |::| :     :|  :
Db        183 DT--FLDHFR---LKHIVKSYSDHIAVPIYFFDEAGNNEIQLNSASALWTRPKSEITEEQ 237

Qy        118 FQYGYKNIRNAIE-------YQVGGGLNITDTLIFTDGELCDVFYVPNADQGCELWVKK- 169
              ::  ||::  ||:       : | : |: |      :  |:|: |:  :  :|::|:
Db        238 YKEFYKSLSYAIDDPWITMHNKNEGAIEFTNLLFIPSSKTFDLFH-PDRKRRVKLYIKRV 296

Qy        170 ----SHYKHVPDYCTFVFNVFCAKD 190
                  :   :| |  |:  |  ::|
Db        297 FISDENIDLIPSYLRFLRGVVDSED 321


RESULT 6
HTPG_BORBU
ID   HTPG_BORBU     STANDARD;      PRT;   616 AA.
AC   P42555;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Chaperone protein htpG (Heat shock protein htpG) (High temperature
DE   protein G).
GN   HTPG OR BB0560.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=297;
RA   Porcella S.F., Radolf J.D., Norgard M.V.;
RL   Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=98065943; PubMed=9403685;
RA   Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA   Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA   Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA   Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA   van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA   Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA   Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA   Smith H.O., Venter J.C.;
RT   "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT   burgdorferi.";
RL   Nature 390:580-586(1997).
RN   [3]
RP   SEQUENCE OF 497-574 FROM N.A.
RC   STRAIN=212;
RX   MEDLINE=95111614; PubMed=7812434;
RA   Ojaimi C., Davidson B.E., Saint-Girons I., Old I.G.;
RT   "Conservation of gene arrangement and an unusual organization of rRNA
RT   genes in the linear chromosomes of the Lyme disease spirochaetes
RT   Borrelia burgdorferi, B. garinii and B. afzelii.";
RL   Microbiology 140:2931-2940(1994).
CC   -!- FUNCTION: Molecular chaperone. Has ATPase activity (By
CC       similarity).
CC   -!- SUBUNIT: Homodimer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Belongs to the heat shock protein 90 family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U51878; AAA97469.1; -.
DR   EMBL; AE001157; AAC66919.1; ALT_INIT.
DR   EMBL; L32145; AAC41403.1; -.
DR   HSSP; P07900; 1YER.
DR   TIGR; BB0560; -.
DR   HAMAP; MF_00505; -; 1.
DR   InterPro; IPR003594; ATPbind_ATPase.
DR   InterPro; IPR001404; Hsp90.
DR   Pfam; PF02518; HATPase_c; 1.
DR   Pfam; PF00183; HSP90; 2.
DR   PRINTS; PR00775; HEATSHOCK90.
DR   SMART; SM00387; HATPase_c; 1.
DR   PROSITE; PS00298; HSP90; 1.
KW   Chaperone; ATP-binding; Heat shock; Complete proteome.
FT   DOMAIN        1    333       A; SUBSTRATE-BINDING (BY SIMILARITY).
FT   DOMAIN      334    542       B (BY SIMILARITY).
FT   DOMAIN      543    616       C.
SQ   SEQUENCE   616 AA;  71218 MW;  D676552F48DBEE84 CRC64;

  Query Match             8.5%;  Score 96;  DB 1;  Length 616;
  Best Local Similarity   21.5%;  Pred. No. 0.69;
  Matches   46;  Conservative   34;  Mismatches   92;  Indels   42;  Gaps    8;

Qy         11 FVSAALATQ-AETTSAKAGENP--LWAHEELLGKYQDAWKSIDQGVSVTYVLAKTTYEND 67
              | || : ::  | || || |:   :|: :   |  :  |  : |  :  | |   |
Db        121 FYSAFIVSEKVEVTSKKALESDAYIWSSDGKTGYEIEKAKKEESGTEIKLYLNKEGLE-- 178

Qy         68 TGSWGSQFKCLQV-------------------------QEIERKEEDYTVTSVFTFRNA 101
                 : :::|  ::                        : || |||    |:   :|
Db        179 ---YANKWKIQEIIKKYSNHINYPIYIKYSEPIMKDGKQEGIEEKEEKLNETTALWTKN- 234

Qy        102 SSPIKYYNVTETVKAVFQYGYKNIRNAIEYQVGGGLNITDTLIFTDGELCDVFYVPNADQ 161
              | ||     |  |   : |:|    |  :  | |  |:          |::| ||
Db        235 KSEIKAEEYNEFYKNT-TFDYENPLMHIHTKAEGNLEYTNLFYVPSKAPYDLYY-PNTKP 292

Qy        162 GCELWVKK-----SHYKHVPDYCTFVFNVFCAKD 190
              | :|:: :     |    :|:|  |:  :  :|
Db        293 GVKLFINRIFITDSEGSLLPNYLRFIKGIIDCQD 326


RESULT 7
HTPG_TREPA
ID   HTPG_TREPA     STANDARD;      PRT;   639 AA.
AC   O83949;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Chaperone protein htpG (Heat shock protein htpG) (High temperature
DE   protein G).
GN   HTPG OR TP0984.
OS   Treponema pallidum.
OC   Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX   NCBI_TaxID=160;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Nichols;
RX   MEDLINE=98332770; PubMed=9665876;
RA   Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
RA   Dodson R., Gwinn M., Hickey E.K., Clayton R., Ketchum K.A.,
RA   Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA   Khalak H., Richardson D., Howell J.K., Chidambaram M., Utterback T.,
RA   McDonald L., Artiach P., Bowman C., Cotton M.D., Fujii C., Garland S.,
RA   Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
RA   Venter J.C.;
RT   "Complete genome sequence of Treponema pallidum, the syphilis
RT   spirochete.";
RL   Science 281:375-388(1998).
CC   -!- FUNCTION: Molecular chaperone. Has ATPase activity (By
CC       similarity).
CC   -!- SUBUNIT: Homodimer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Belongs to the heat shock protein 90 family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
```

```
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ------------------------------------------------------------------------
DR   EMBL; AE001265; AAC65938.1; -.
DR   PIR; F71258; F71258.
DR   HSSP; P02829; 1AH8.
DR   TIGR; TP0984; -.
DR   HAMAP; MF_00505; -; 1.
DR   InterPro; IPR003594; ATPbind_ATPase.
DR   InterPro; IPR001404; Hsp90.
DR   Pfam; PF02518; HATPase_c; 1.
DR   Pfam; PF00183; HSP90; 2.
DR   PRINTS; PR00775; HEATSHOCK90.
DR   SMART; SM00387; HATPase_c; 1.
DR   PROSITE; PS00298; HSP90; 1.
KW   Chaperone; ATP-binding; Heat shock; Complete proteome.
FT   DOMAIN        1    348       A; SUBSTRATE-BINDING (BY SIMILARITY).
FT   DOMAIN      349    565       B (BY SIMILARITY).
FT   DOMAIN      566    639       C.
SQ   SEQUENCE   639 AA;  72937 MW;  3E8FDBAC2282C31D CRC64;

  Query Match             7.7%;  Score 87;  DB 1;  Length 639;
  Best Local Similarity   22.9%;  Pred. No. 4.5;
  Matches   50;  Conservative   26;  Mismatches   52;  Indels   90;  Gaps   10;

Qy          5 VVLLLTFVSAALATQAETTSAKAGENPLWAHEELLGKYQD----------AWKSIDQGVS 54
              ||| |:  ::  ||:           |  ||:: || |           |  |:  :
Db        182 VVLHLSQENSEFATR------------WRLEEVIKKYSDHIAFPIYLHYLQKEYDKDGA 228

Qy         55 VTYVLAKTTYENDTGSWGSQFKCLQVQEIERKEEDY-----TVTSVFTFRNASSPIKYYN 109
              ||    |  || |:   : |      | |||||     |:|     |::: |: | :
Db        229 VTDTQKKVDQVNDAGALWKRPKS------ELKEEDYHRFYQTLT-----RDSTPPLLYVH 277

Qy        110 VTETVKAVFQYGYKNIRNAIEYQVGGGLNITDTLIFTDG--ELCDVFYVP--------NA 159
                                                :|  |   :||||        :|
Db        278 TK---------------------------------AEGTQEYVTLFYVPAKAPFDLFHA 303

Qy        160 D--QGCELWVKK-----SHYKHVPDYCTFVFNVFCAKD 190
              |   | :|:||:        : :| |  ||  |  ::|
Db        304 DYKPGVKLFVKRVFITDDEKELLPVYLRFVRGVIDSED 341


RESULT 8
FAF_DROME
ID   FAF_DROME      STANDARD;      PRT;  2778 AA.
AC   P55824; Q9V9T6; Q9Y027;
DT   01-NOV-1997 (Rel. 35, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Probable ubiquitin carboxyl-terminal hydrolase FAF (EC 3.1.2.15)
DE   (Ubiquitin thiolesterase FAF) (Ubiquitin-specific processing protease
DE   FAF) (Deubiquitinating enzyme FAF) (Fat facets protein).
GN   FAF OR BCDNA:LD22582 OR CG1945.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORMS 2 AND 3), AND TISSUE SPECIFICITY.
RC   TISSUE=Eye imaginal disk;
RX   MEDLINE=93202020; PubMed=1295747;
RA   Fischer-Vize J.A., Rubin G.M., Lehmann R.;
RT   "The fat facets gene is required for Drosophila eye and embryo
RT   development.";
RL   Development 116:985-1000(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [3]
RP   REVISIONS, AND ALTERNATIVE SPLICING.
RX   MEDLINE=22426069; PubMed=12537572;
RA   Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA   Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA   Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA   Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA   Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA   Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA   Lewis S.E.;
RT   "Annotation of the Drosophila melanogaster euchromatic genome: a
RT   systematic review.";
RL   Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN   [4]
RP   SEQUENCE OF 1089-2778 FROM N.A. (ISOFORM 1).
RC   STRAIN=Berkeley;
RX   MEDLINE=20196012; PubMed=10731138;
RA   Rubin G.M., Hong L., Brokstein P., Evans-Holm M., Frise E.,
RA   Stapleton M., Harvey D.A.;
RT   "A Drosophila complementary DNA resource.";
RL   Science 287:2222-2224(2000).
CC   -!- FUNCTION: Required for eye and embryo development, and plays a
CC       role in compound eye assembly and oogenesis respectively. In the
CC       larval eye disks, cells outside the assembling facets require this
CC       protein for short-range cell interactions that prevent the mystery
CC       cells from becoming photoreceptors. It is also required for
CC       nuclear migration and cellularization in early embryogenesis and
CC       could play a role in pole cell determination, development or
CC       function.
CC   -!- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O =
CC       ubiquitin + a thiol.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC         Comment=Experimental confirmation may be lacking for some
CC         isoforms;
CC       Name=1;
CC         IsoId=P55824-1; Sequence=Displayed;
CC       Name=2;
```

```
CC         IsoId=P55824-2; Sequence=VSP_005270;
CC       Name=3;
CC         IsoId=P55824-3; Sequence=VSP_005269;
CC   -!- TISSUE SPECIFICITY: Eye disks and ovaries.
CC   -!- DEVELOPMENTAL STAGE: Expressed both maternally and zygotically.
CC   -!- SIMILARITY: Belongs to peptidase family C19.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L04959; AAF01345.1; -.
DR   EMBL; L04958; AAF01346.1; -.
DR   EMBL; L04960; AAF01347.1; -.
DR   EMBL; L04960; AAF01348.1; -.
DR   EMBL; AE003779; AAF57198.1; -.
DR   EMBL; AE003779; AAN14291.1; -.
DR   EMBL; AF145677; AAD38652.1; -.
DR   MEROPS; C19.007; -.
DR   FlyBase; FBgn0005632; faf.
DR   GO; GO:0005737; C:cytoplasm; IDA.
DR   GO; GO:0007349; P:cellularization; IMP.
DR   GO; GO:0009795; P:embryonic morphogenesis; IMP.
DR   GO; GO:0007456; P:eye morphogenesis (sensu Drosophila); IMP.
DR   GO; GO:0008583; P:mystery cell fate differentiation (sensu Dr. . .; IMP.
DR   GO; GO:0007097; P:nuclear migration; IMP.
DR   GO; GO:0016579; P:protein deubiquitination; IDA.
DR   GO; GO:0006511; P:ubiquitin-dependent protein catabolism; IGI.
DR   InterPro; IPR001394; Peptidase_C19.
DR   Pfam; PF00443; UCH; 1.
DR   PROSITE; PS00972; UCH_2_1; 1.
DR   PROSITE; PS00973; UCH_2_2; 1.
DR   PROSITE; PS50235; UCH_2_3; 1.
KW   Ubl conjugation pathway; Hydrolase; Thiol protease;
KW   Developmental protein; Vision; Alternative splicing.
FT   ACT_SITE   1677   1677       BY SIMILARITY.
FT   ACT_SITE   1978   1978       BY SIMILARITY.
FT   ACT_SITE   1986   1986       BY SIMILARITY.
FT   VARSPLIC   2705   2778       KCRRVIIKKLVESKDEEDATTATTAATTEVTTSPATAIATA
FT                                ATLEPAGMSELTTMVEKNLIISQENPQAKSSLQ -> VTRA
FT                                NNV (in isoform 3).
FT                                /FTId=VSP_005269.
FT   VARSPLIC   2742   2778       IATAATLEPAGMSELTTMVEKNLIISQENPQAKSSLQ ->
FT                                SQRQQL (in isoform 2).
FT                                /FTId=VSP_005270.
FT   CONFLICT    234    234       E -> D (IN REF. 1).
FT   CONFLICT   2725   2725       T -> S (IN REF. 1; AAF01345).
SQ   SEQUENCE   2778 AA;  311139 MW;  FFB90438BA53A02B CRC64;

  Query Match             7.6%;  Score 86;  DB 1;  Length 2778;
  Best Local Similarity   22.9%;  Pred. No. 33;
  Matches   54;  Conservative   28;  Mismatches   72;  Indels   82;  Gaps   12;

Qy          6 VLLLTFVSAALATQAETTSAKAGENPLWAHEELLG------KYQDA-------WKSIDQG 52
              |::|  | |  |    :          || | :|||      : |||        :|:|:|
Db       1756 VVILKHVQAIFAHLGHSALQYYVPRGLWTHFKLLGEPVNLREQQDAVEFFMSLLESLDEG 1815

Qy         53 VSVTYVLAKTTYENDT--GSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASSPIKYYNV 110
              :    | :    | |  ||: |  | :      ||| ::| ||   || ||      :
Db       1816 LK---ALGQPQLMNATLGGSFSDQKICQECPHRYSKEEPFSVFSV-DIRNHSS------L 1865

Qy        111 TETVKAVFQYGYKNIRNAIEYQVGGGLNITDTLIFTDGEL---CDVFYVPNADQGCELWV 167
              ||:::                          :  |||    | ::    |:  : |
Db       1866 TESLEQ--------------------------YVKGELLEGADAYHCDKCDKKV-VTV 1896

Qy        168 KKSHYKHVP-----------DY---CTFVFNVFCAKDRKTYDIFNEECVYNGEPW 208
              |:   | :|                ||   |   ||          | |    : : ||:
Db       1897 KRVCVKKLPPVLAIQLKRFEYDYERVCAIKFN----------DYFEPPRILDMEPY 1942
```

```
RESULT 9
HTPG_MYCLE
ID   HTPG_MYCLE     STANDARD;      PRT;   656 AA.
AC   O33012;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Chaperone protein htpG (Heat shock protein htpG) (High temperature
DE   protein G).
GN   HTPG OR ML1623 OR MLCB250.19C.
OS   Mycobacterium leprae.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1769;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TN;
RX   MEDLINE=21128732; PubMed=11234002;
RA   Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA   Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA   Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA   Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA   Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA   Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA   Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA   Barrell B.G.;
RT   "Massive gene decay in the leprosy bacillus.";
RL   Nature 409:1007-1011(2001).
CC   -!- FUNCTION: Molecular chaperone. Has ATPase activity (By
CC       similarity).
CC   -!- SUBUNIT: Homodimer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Belongs to the heat shock protein 90 family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z97369; CAB10613.1; -.
DR   EMBL; AL583922; CAC30574.1; -.
DR   PIR; A87112; A87112.
DR   HSSP; P07900; 1BYQ.
DR   Leproma; ML1623; -.
DR   HAMAP; MF_00505; -; 1.
DR   InterPro; IPR003594; ATPbind_ATPase.
DR   InterPro; IPR001404; Hsp90.
DR   Pfam; PF02518; HATPase_c; 1.
DR   Pfam; PF00183; HSP90; 2.
DR   PRINTS; PR00775; HEATSHOCK90.
DR   SMART; SM00387; HATPase_c; 1.
DR   PROSITE; PS00298; HSP90; 1.
KW   Chaperone; ATP-binding; Heat shock; Complete proteome.
FT   DOMAIN        1    359       A; SUBSTRATE-BINDING (BY SIMILARITY).
FT   DOMAIN      360    575       B (BY SIMILARITY).
FT   DOMAIN      576    656       C.
SQ   SEQUENCE   656 AA;  73866 MW;  A526690CA66E03FF CRC64;

  Query Match             7.6%;  Score 85;  DB 1;  Length 656;
  Best Local Similarity   21.3%;  Pred. No. 7;
  Matches   49;  Conservative   38;  Mismatches   79;  Indels   64;  Gaps    9;

Qy         11 FVSAALATQAETTSAKAGENPL--WAHEELLGKYQDAWKSID---QGVSVTYVLAKTTYE 65
              : |  :| : |  : ||||      |: :   |:       :|:|    || |||    | :|
Db        137 YSSFMVANKVELLTRKAGETAATRWSSD---GEATYTIESVDEAPQGTSVTLHLKPEDFE 193

Qy         66 NDTGSWGSQFKCLQV-------------QEIERK------------EEDYTVTSVFTFRN 100
              ::   : |::|  ::              |:||:            ||  |: :     |
```

```
Db        194 DELHDYTSEWKIRELVKKYSDFIAWPIRMEVERRAPATSDGEGADGEEQVTIETQTI--N 251

Qy        101 ASSPIKYYNVTETVKAVFQYGYKNIRNA---------------IEYQVGGGLNITDTLIF 145
              :   :   :  |  :  ::  ||:| :|               |||          |:|
Db        252 SMKALWTKSKDEVSEDEYKEFYKHIAHAWDDPLEVIAMKAEGTFEYQ---------ALLF 302

Qy        146 TDGELCDVFYVPNADQGCELWVKKSHY-----KHVPDYCTFVFNVFCAKD 190
                        :  :|  | :|:||:       : :| |  ||  |  |:|
Db        303 IPSHAPFDLFNSDAKIGMQLYVKRVFIMSDCDQLMPMYLRFVKGVVDAED 352


RESULT 10
HTPG_RICPR
ID   HTPG_RICPR     STANDARD;      PRT;   621 AA.
AC   Q9ZCB9;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Chaperone protein htpG (Heat shock protein htpG) (High temperature
DE   protein G).
GN   HTPG OR RP840.
OS   Rickettsia prowazekii.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=782;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Madrid E;
RX   MEDLINE=99039499; PubMed=9823893;
RA   Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA   Sicheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA   Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT   "The genome sequence of Rickettsia prowazekii and the origin of
RT   mitochondria.";
RL   Nature 396:133-140(1998).
CC   -!- FUNCTION: Molecular chaperone. Has ATPase activity (By
CC       similarity).
CC   -!- SUBUNIT: Homodimer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Belongs to the heat shock protein 90 family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AJ235273; CAA15264.1; -.
DR   PIR; H71645; H71645.
DR   HSSP; P02829; 1A4H.
DR   HAMAP; MF_00505; -; 1.
DR   InterPro; IPR003594; ATPbind_ATPase.
DR   InterPro; IPR001404; Hsp90.
DR   Pfam; PF02518; HATPase_c; 1.
DR   Pfam; PF00183; HSP90; 2.
DR   PRINTS; PR00775; HEATSHOCK90.
DR   SMART; SM00387; HATPase_c; 1.
DR   PROSITE; PS00298; HSP90; 1.
KW   Chaperone; ATP-binding; Heat shock; Complete proteome.
FT   DOMAIN       1    328       A; SUBSTRATE-BINDING (BY SIMILARITY).
FT   DOMAIN     329    544       B (BY SIMILARITY).
FT   DOMAIN     545    621       C.
SQ   SEQUENCE   621 AA;  70713 MW;  9F97B97801524007 CRC64;

  Query Match             7.5%;  Score 84.5;  DB 1;  Length 621;
  Best Local Similarity   18.9%;  Pred. No. 7.3;
  Matches   39;  Conservative   42;  Mismatches   88;  Indels   37;  Gaps    8;

Qy         11 FVSAALATQAETTSAKAGENPLWAHE-ELLGKY--QDAWKSIDQGVSVTYVLAK--TTYE 65
              :  |  :| :   || ||||: :   | : ||:|   |: :   :|  :   : |  ||:
Db        127 YSSFMVADKVTVTSRKAGESKVHTWESDGLGEYIVADSEQEFTRGTEIVLYIKKSETTFL 186

Qy         66 NDTGSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASSPIKYYNV-----TETVKAVFQY 120
              :             ::: | :  |:    ::    | :      |      |    : :
Db        187 DH----------FRLKHIVKSYSDHIAVPIYFCDEAGNNEIQLNSASALWTRPKSEITED 236

Qy        121 GYKNIRNAIEYQVG----------GGLNITDTLIFTDGELCDVFYVPNADQGCELWVKK 169
               ||    :: | :|           | :  |: |     :  |:|: |:  :  :|::|:
Db        237 QYKEFYKSLSYAVDDPWVTLHNKNEGAIEFTNLLFIPSSKTFDLFH-PDRKKRVKLYIKR 295

Qy        170 -----SHYKHVPDYCTFVFNVFCAKD 190
                    :   :| |  |:  |  ::|
Db        296 VFISDENIDLIPSYLRFLRGVVDSED 321


RESULT 11
Y692_METTH
ID   Y692_METTH     STANDARD;      PRT;   318 AA.
AC   O26788;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Hypothetical protein MTH692.
GN   MTH692.
OS   Methanobacterium thermoautotrophicum.
OC   Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC   Methanobacteriaceae; Methanothermobacter.
OX   NCBI_TaxID=187420;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Delta H;
RX   MEDLINE=98037514; PubMed=9371463;
RA   Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA   Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA   Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA   Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA   Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA   McDougall S., Shimer G., Goyal A., Pietrovski S., Church G.M.,
RA   Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT   "Complete genome sequence of Methanobacterium thermoautotrophicum
RT   deltaH: functional analysis and comparative genomics.";
RL   J. Bacteriol. 179:7135-7155(1997).
CC   -!- SIMILARITY: Belongs to the band 7 / mec-2 family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AE000848; AAB85197.1; -.
DR   PIR; C69192; C69192.
DR   InterPro; IPR001107; Band_7.
DR   InterPro; IPR001972; Stomatin.
DR   Pfam; PF01145; Band_7; 1.
DR   PRINTS; PR00721; STOMATIN.
DR   SMART; SM00244; PHB; 1.
DR   PROSITE; PS01270; BAND_7; 1.
KW   Hypothetical protein; Transmembrane; Complete proteome.
FT   TRANSMEM     2     22       POTENTIAL.
SQ   SEQUENCE   318 AA;  35425 MW;  4A72C0AC8E99278D CRC64;

  Query Match             7.3%;  Score 82.5;  DB 1;  Length 318;
  Best Local Similarity   23.3%;  Pred. No. 4.9;
  Matches   38;  Conservative   23;  Mismatches   77;  Indels   25;  Gaps    6;

Qy          5 VVLLLTFVSAALATQAETTSAKAGENPLWAHEELLGKYQDAWKSIDQGVSVTYVLAKTTY 64
              |:|:| | |  :   |                 | |||||   :::: |: |     :
Db         10 VLLVLAFKSLKILRPYEKGVV----------ERLGKYQ---RTVESGLVVIIPFIEAIK 55

Qy         65 ENDTGSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASSPIK-YYNVTETVKAVFQYGYK 123
              : |       |   :  ||:  |:    |     |    |    |||  :  :|: :
Db         56 ENDTGSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASSPIKYYNVTETVKAVFQYGYK 115
```

```
Db         56 KVD---MREQVVDVPPQEVITKDNTVVVVDCVIFYEVVDPFNAVYNVVDFYQAITKLAQT 112

Qy        124 NIRNAIEYQVGGGLNITDTLIFTDGELCDVFYVPNADQGCELW 166
              |:|| |     | | :  ||  |  |: :       |:  : |
Db        113 NLRNII-----GDLELDQTL--TSREMINTQLREVLDEATDKW 148


RESULT 12
RIR2_CAEEL
ID   RIR2_CAEEL     STANDARD;      PRT;   381 AA.
AC   P42170;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Ribonucleoside-diphosphate reductase small chain (EC 1.17.4.1)
DE   (Ribonucleotide reductase).
GN   RNR-2 OR C03C10.3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Bristol N2;
RA   Berks M.;
RL   Submitted (AUG-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Provides the precursors necessary for DNA synthesis.
CC   -!- CATALYTIC ACTIVITY: 2'-deoxyribonucleoside diphosphate + oxidized
CC       thioredoxin + H(2)O = ribonucleoside diphosphate + reduced
CC       thioredoxin.
CC   -!- COFACTOR: Binds 2 iron ions per subunit (By similarity).
CC   -!- PATHWAY: DNA replication pathway; first step.
CC   -!- SUBUNIT: Heterodimer of a large and a small chain.
CC   -!- SIMILARITY: Belongs to the ribonucleoside diphosphate reductase
CC       small chain family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z35637; CAA84688.1; -.
DR   PIR; T18876; T18876.
DR   HSSP; P11157; 1XSM.
DR   WormPep; C03C10.3; CE00874.
DR   InterPro; IPR000358; Ribonucl_redctse.
DR   Pfam; PF00268; ribonuc_red_sm; 1.
DR   PROSITE; PS00368; RIBORED_SMALL; 1.
KW   Oxidoreductase; DNA replication; Metal-binding; Iron.
FT   METAL       130    130       IRON 1 (BY SIMILARITY).
FT   METAL       161    161       IRON 1 AND 2 (BY SIMILARITY).
FT   METAL       164    164       IRON 1 (BY SIMILARITY).*
FT   METAL       224    224       IRON 2 (BY SIMILARITY).
FT   METAL       258    258       IRON 2 (BY SIMILARITY).
FT   METAL       261    261       IRON 2 (BY SIMILARITY).
FT   ACT_SITE    168    168       BY SIMILARITY.
SQ   SEQUENCE   381 AA;  44289 MW;  75497147ABF36C59 CRC64;

  Query Match             7.3%;  Score 82;  DB 1;  Length 381;
  Best Local Similarity   21.8%;  Pred. No. 6.7;
  Matches   29;  Conservative   20;  Mismatches   56;  Indels   28;  Gaps    3;

Qy          9 LTFVSAALATQAETTSAKAGENPLWAHEELLG----------KYQDAWKSIDQGVSVTY 57
              |  |     |  || |: ::  | | | | :|          |: | |    : |:  :
Db         35 LEIVDQTKAASAEETNNESEVNELDADEPMLQDLDNRFVIFPLKHHDIWNFYKKAVASFW 94

Qy         58 VLAKTTYENDTGSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASSPI-------KYYNV 110
               : :      |   |            |    :| | :: :  |  || .|       :: |
Db         95 TVEEVDLGKDMNDW----------EKMNGDEQYFISRILAFFAASDGIVNENLCERFSNE 144

Qy        111 TETVKAVFQYGYK 123
                :  :| | ||::
Db        145 VQVSEARFFYGFQ 157


RESULT 13
TRG5_ECOLI
ID   TRG5_ECOLI     STANDARD;      PRT;   637 AA.
AC   Q00184;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Conjugal transfer protein traG.
GN   TRAG.
OS   Escherichia coli.
OG   Plasmid IncP-beta R751.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HB101;
RX   MEDLINE=92190548; PubMed=1665997;
RA   Ziegelin G., Pansegrau W., Strack B., Balzer D., Kroeger M.,
RA   Kruft V., Lanka E.;
RT   "Nucleotide sequence and organization of genes flanking the transfer
RT   origin of promiscuous plasmid RP4.";
RL   DNA Seq. 1:303-327(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Thomas C.M.;
RL   Submitted (AUG-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Required for conjugative transfert of plasmid R751.
CC       Binds tightly and specifically to the relaxase traI. Can also bind
CC       to DNA without sequence specificity. May form a pore-like
CC       structure that could serve as a channel for DNA transfer (By
CC       similarity).
CC   -!- SUBUNIT: May form multimers of at least 18 subunits (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC       (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE VIRD4/TRAG FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X54458; CAA38327.1; -.
DR   EMBL; U67194; AAC64474.1; -.
DR   InterPro; IPR003688; TRAG.
DR   Pfam; PF02534; TRAG; 1.
KW   Plasmid; Conjugation; DNA-binding; Transmembrane; Inner membrane.
FT   DOMAIN        1     22       CYTOPLASMIC (POTENTIAL).
FT   TRANSMEM     23     43       POTENTIAL.
FT   DOMAIN       44     84       PERIPLASMIC (POTENTIAL).
FT   TRANSMEM     85    105       POTENTIAL.
FT   DOMAIN      106    637       CYTOPLASMIC (POTENTIAL).
SQ   SEQUENCE   637 AA;  69883 MW;  7B45A9A9020902FB CRC64;

  Query Match             7.3%;  Score 82;  DB 1;  Length 637;
  Best Local Similarity   18.3%;  Pred. No. 12;
  Matches   46;  Conservative   30;  Mismatches   89;  Indels   86;  Gaps    7;

Qy          5 VVLLLTFVSAALATQAETTSAKAGE----NPLWAHEE---------------------- 37
              :|  :  :  |:|    : |:|| |    :  || ::
Db         90 LVSTVGLLGVAVAKVVTSNSSKANEYLHGSARWAEKKDIQAAGLLPRERNVLEIVTGKAA 149

Qy         38 --LLGKYQDAWKSID--------QGVSVTYVLAKTTYENDTG-------SWGSQFKCLQV 80
                 | |   |:  |          |       | |       |       |||:      :
```

```
Db        150 PTATGVYVGGWQDKDGNFFYLRHSGPEHVLTYAPTRSGKGVGLVVPTLLSWGASSVITDL 209

Qy         81 QEIERKEEDYTVTSVFTFRNASSPIKYYNVTETVKAVFQYGYKNIRNAIEYQVGGGLNIT 140
                 | | : :|: :  ::| : :  :     |    |         ||   ||:||   |:
Db        210 -----KGELWALTAGWRQKHAKNKVLRFEPASTSGGVCWNPLDEIRLGTEYEVGDVQNLA 264

Qy        141 DTLIFTDGELCDVFYV----------------------------------PNADQGCE 164
                 ::  ||:  |  :                                  || | | |
Db        265 TLIVDPDGKGLDSHWQKTAFALLVGVILHALYKAKDDGGTATLPSVDAMLADPNRDIG-E 323

Qy        165 LWVKKSHYKHV 175
                ||:: : | ||
Db        324 LWMEMATYGHV 334


RESULT 14
DNAK_RICPR
ID   DNAK_RICPR     STANDARD;      PRT;   627 AA.
AC   Q9ZDX9;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Chaperone protein dnaK (Heat shock protein 70) (Heat shock 70 kDa
DE   protein) (HSP70).
GN   DNAK OR RP185.
OS   Rickettsia prowazekii.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rickettsiales;
OC   Rickettsiaceae; Rickettsieae; Rickettsia.
OX   NCBI_TaxID=782;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Madrid E;
RX   MEDLINE=99039499; PubMed=9823893;
RA   Andersson S.G.E., Zomorodipour A., Andersson J.O.,
RA   Sicheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
RA   Eriksson A.-S., Winkler H.H., Kurland C.G.;
RT   "The genome sequence of Rickettsia prowazekii and the origin of
RT   mitochondria.";
RL   Nature 396:133-140(1998).
CC   -!- FUNCTION: Acts as a chaperone (By similarity).
CC   -!- INDUCTION: By stress conditions e.g. heat shock (By similarity).
CC   -!- SIMILARITY: Belongs to the heat shock protein 70 family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AJ235270; CAA14651.1; -.
DR   PIR; D71729; D71729.
DR   HSSP; P04475; 1DG4.
DR   HAMAP; MF_00332; -; 1.
DR   InterPro; IPR001023; Hsp70.
DR   Pfam; PF00012; HSP70; 1.
DR   PRINTS; PR00301; HEATSHOCK70.
DR   ProDom; PD000089; Hsp70; 1.
DR   PROSITE; PS00297; HSP70_1; 1.
DR   PROSITE; PS00329; HSP70_2; 1.
DR   PROSITE; PS01036; HSP70_3; 1.
KW   Chaperone; ATP-binding; Heat shock; Phosphorylation;
KW   Complete proteome.
FT   MOD_RES     197    197       PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ   SEQUENCE   627 AA;  68383 MW;  A638A0A5A5C8ACC7 CRC64;

  Query Match             7.2%;  Score 81.5;  DB 1;  Length 627;
  Best Local Similarity   28.2%;  Pred. No. 14;
  Matches   33;  Conservative   15;  Mismatches   42;  Indels   27;  Gaps    7;

Qy         72 GSQFKCLQVQE------IERKEEDYTVTSVFTFRNA----SSPIKYYNVTETVKAVFQYG 121
              |:   |: | |       |:  | : |  |:  | |:      | |   ||

Db         10 GTTNSCVAVMEGKEPKVIDNAEGERTTPSIIAFANSERLVGQPAKRQAVT---------- 59

Qy        122 YKNIRNAIEYQVGG--GLNITDTLIFTDGELCDVFYVPNADQGCELWVKKSHYKHVP 176
                | || | | |    | | ||  ::  |  |   : :  || | : ||:  ::|: |
Db         60 --NPRNTI-YAVKRLIGRNFTDPMVRKDQGLVP-YNIVKADNG-DAWVEADNHKYSP 111


RESULT 15
HS9C_DICDI
ID   HS9C_DICDI     STANDARD;      PRT;   700 AA.
AC   P54651;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Heat shock cognate 90 kDa protein.
GN   HSPD OR HSC90.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AX3;
RA   Boves H., Dittrich W., Mintert U., Lottspeich F., Gerisch G.,
RA   Faix J.;
RL   Submitted (JUL-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Molecular chaperone. Has ATPase activity (By
CC       similarity).
CC   -!- SIMILARITY: Belongs to the heat shock protein 90 family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L43591; AAA69917.1; -.
DR   HSSP; P07900; 1YER.
DR   DictyBase; DDB0001943; hspD.
DR   InterPro; IPR003594; ATPbind_ATPase.
DR   InterPro; IPR001404; Hsp90.
DR   Pfam; PF02518; HATPase_c; 1.
DR   Pfam; PF00183; HSP90; 1.
DR   PRINTS; PR00775; HEATSHOCK90.
DR   SMART; SM00387; HATPase_c; 1.
DR   PROSITE; PS00298; HSP90; 1.
KW   Chaperone; ATP-binding; Heat shock.
SQ   SEQUENCE   700 AA;  79881 MW;  BFB5ECF238089CD3 CRC64;

  Query Match             7.2%;  Score 81;  DB 1;  Length 700;
  Best Local Similarity   19.5%;  Pred. No. 17;
  Matches   52;  Conservative   32;  Mismatches  105;  Indels   78;  Gaps    9;

Qy         11 FVSAALATQAETTSAKAG--ENPLWA-----------HEELLGK--------------Y 42
              | ||| |          :|    |  :|           | | ||:             |
Db        128 FYSAYLVADTVIVHSKNNDDEQYVWESSAGGEFTIALDHTEPLGRGTKIVLHMKEDQLDY 187

Qy         43 QDAWK----------SIDQGVSVTYVLAKTTYENDTGSWGSQFKC-LQVQEIERKEEDYT 91
              | |                |  :|:  :  |   |  |  |   | :     :::||| ::|
Db        188 LDETKIKNLVKKHSEFIQYPISLLTIKEKEVDEETTAKEGEEESTDAKIEEIEEEKEKKK 247

Qy         92 V---TSVFTFRNASSPIKYYNVTETVKAVFQYGYKNIRNAIE-------YQVGGGLNITD 141
              |       :   | : |:   | ::  |  :   ||:| |  |        : | | |
Db        248 VKVQEKEWDVLNKTKPLWTRNPSDVTKEEYNSFYKSISNDWEEPLAVKHFSVEGQLEFKA 307

Qy        142 TLIFTDGELCDVFYVPNADQGCELWVKKSHYKH-----VPDYCTFVFNVFCAKD------ 190
              |         |:|          :|:||:          :|:|  ||  :  ::|
Db        308 ILFVPKKAPFDLFESKKKANNIKLYVKRVFIMDNCADIIPEYLNFVRGIVDSEDLPLNIS 367

Qy        191 ------------------RKTYDIFNE 199
                               :|  ::|||
Db        368 RETLQQNKILTVIRKNLVKKCIELFNE 394
```

Search completed: March 9, 2004, 13:34:32
Job time : 20 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 9, 2004, 13:31:31 ; Search time 46 Seconds
(without alignments)
1433.550 Million cell updates/sec

Title: US-09-555-296B-4
Perfect score: 1125
Sequence: 1 MKMQVVLLLTFVSAALATQA.....DRKTYDIFNEECVYNGEPWL 209

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1125	100.0	209	5	Q8WSK7	Q8wsk7 dermacentor
2	126	11.2	210	5	Q8MV98	Q8mv98 ixodes scap
3	117.5	10.4	306	5	Q8MVC3	Q8mvc3 ixodes scap
4	111.5	9.9	176	5	Q8I9T9	Q8i9t9 ornithodoro
5	103	9.2	311	5	Q8MVC2	Q8mvc2 ixodes scap
6	102.5	9.1	722	16	Q837S0	Q837s0 enterococcu
7	97.5	8.7	221	5	Q95WZ5	Q95wz5 ixodes scap
8	95	8.4	508	16	Q8A286	Q8a286 bacteroides
9	89	7.9	972	5	Q8IDR7	Q8idr7 plasmodium
10	87	7.7	598	2	Q939L2	Q939l2 citrobacter
11	86.5	7.7	932	16	Q92N47	Q92n47 rhizobium m
12	86	7.6	399	16	Q8XMD0	Q8xmd0 clostridium
13	86	7.6	584	2	Q51717	Q51717 pseudomonas
14	85.5	7.6	917	2	O33983	O33983 listeria mo
15	85.5	7.6	1090	16	Q8A8I2	Q8a8i2 bacteroides
16	85.5	7.6	6761	5	Q8IC77	Q8ic77 plasmodium
17	85	7.6	437	10	Q9SQX6	Q9sqx6 arabidopsis
18	85	7.6	476	17	Q8ZWE9	Q8zwe9 pyrobaculum
19	84.5	7.5	524	12	Q9EN29	Q9en29 amsacta moo
20	84.5	7.5	917	2	O50272	O50272 listeria mo
21	84.5	7.5	917	16	Q8Y496	Q8y496 listeria mo
22	84.5	7.5	1801	5	Q8WSJ2	Q8wsj2 bombyx mori
23	84	7.5	218	5	Q95WY8	Q95wy8 ixodes scap
24	84	7.5	386	13	Q8AVY2	Q8avy2 xenopus lae
25	84	7.5	1047	16	Q8A5Y1	Q8a5y1 bacteroides
26	84	7.5	2470	5	Q7YYJ2	Q7yyj2 cryptospori
27	83.5	7.4	364	16	Q8A5I5	Q8a5i5 bacteroides
28	83.5	7.4	444	5	O02158	O02158 caenorhabdi
29	83.5	7.4	1133	12	Q8JRX7	Q8jrx7 phthorimaea
30	83.5	7.4	2879	5	Q8IC57	Q8ic57 plasmodium
31	83	7.4	510	16	Q838U9	Q838u9 enterococcu
32	83	7.4	644	16	Q8FN80	Q8fn80 corynebacte
33	83	7.4	2285	9	O64046	O64046 bacteriopha
34	83	7.4	2285	16	O31976	O31976 bacillus su
35	82.5	7.3	218	5	Q7YT12	Q7yt12 rhodnius pr
36	82.5	7.3	309	16	Q82SA1	Q82sa1 nitrosomona
37	82.5	7.3	324	16	Q898B0	Q898b0 clostridium
38	82	7.3	163	16	Q97DZ8	Q97dz8 clostridium
39	82	7.3	360	12	Q98544	Q98544 paramecium
40	82	7.3	390	16	Q9PQU5	Q9pqu5 ureaplasma
41	82	7.3	561	5	Q9N4D7	Q9n4d7 caenorhabdi
42	82	7.3	564	5	Q95Y41	Q95y41 caenorhabdi
43	82	7.3	637	2	Q937B8	Q937b8 pseudomonas
44	82	7.3	637	2	Q7X3D2	Q7x3d2 gamma-prote
45	82	7.3	713	2	Q9JN39	Q9jn39 mycoplasma

ALIGNMENTS

RESULT 1
Q8WSK7
ID Q8WSK7 PRELIMINARY; PRT; 209 AA.
AC Q8WSK7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Serotonin and histamine binding protein.
OS Dermacentor reticulatus.
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC Parasitiformes; Ixodida; Ixodidae; Dermacentor.
OX NCBI_TaxID=57047;
RN [1]
RP SEQUENCE FROM N.A.
RA Sangamnadech S., Paesen G.C., Nuttall P.A.;
RT "A high affinity serotonin- and histamine-binding lipocalin secreted
RT by blood-feeding ticks.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF217101; AAL56644.1; -.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR002970; His_binding.
DR Pfam; PF02098; His_binding; 1.
DR ProDom; PD152455; His_binding; 1.
SQ SEQUENCE 209 AA; 23889 MW; E5EB8BE710859591 CRC64;

Query Match 100.0%; Score 1125; DB 5; Length 209;
Best Local Similarity 100.0%; Pred. No. 2.8e-92;
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MKMQVVLLLTFVSAALATQAETTSAKAGENPLWAHEELLGKYQDAWKSIDQGVSVTYVLA 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MKMQVVLLLTFVSAALATQAETTSAKAGENPLWAHEELLGKYQDAWKSIDQGVSVTYVLA 60

Qy         61 KTTYENDTGSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASSPIKYYNVTETVKAVFQY 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 KTTYENDTGSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASSPIKYYNVTETVKAVFQY 120

Qy        121 GYKNIRNAIEYQVGGGLNITDTLIFTDGELCDVFYVPNADQGCELWVKKSHYKHVPDYCT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        121 GYKNIRNAIEYQVGGGLNITDTLIFTDGELCDVFYVPNADQGCELWVKKSHYKHVPDYCT 180

Qy        181 FVFNVFCAKDRKTYDIFNEECVYNGEPWL 209
              |||||||||||||||||||||||||||||
Db        181 FVFNVFCAKDRKTYDIFNEECVYNGEPWL 209


RESULT 2
Q8MV98
ID   Q8MV98      PRELIMINARY;      PRT;   210 AA.
AC   Q8MV98;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Putative 22.5 kDa secreted protein.
OS   Ixodes scapularis (Black-legged tick) (Deer tick).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC   Parasitiformes; Ixodida; Ixodidae; Ixodes.
OX   NCBI_TaxID=6945;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Rhode Island; TISSUE=Salivary gland;
RA   Valenzuela J.G., Francischetti I.M., Pham V.M., Garfield M.,
RA   Mather T.N., Ribeiro J.M.C.;
RT   "Exploring the Sialome of the Tick Vector of Lyme Disease, Ixodes
RT   scapularis.";
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF483742; AAM93664.1; -.
SQ   SEQUENCE   210 AA;  24689 MW;  C5B75614649A9BED CRC64;

  Query Match             11.2%;  Score 126;  DB 5;  Length 210;
  Best Local Similarity   24.3%;  Pred. No. 0.0029;
  Matches   53;  Conservative   41;  Mismatches   90;  Indels   34;  Gaps   12;

Qy          6 VLLLTFVSAALATQAETTSAKAGENPLWAHEELLGKYQDAWKSIDQGVSVTYVLAKTTYE 65
              |:  :|:   ||   :   || ||  :   :|   :|||  ::::     :::  : ||
Db          3 VMHFSFLLCLLAILVD---AKPGEIRI-DEDEKYWQYQDIQRALNNPDRESWLYYR-TYR 57

Qy         66 NDTGSWGSQFKCL--QVQEIERKEEDYTVTSVFTFRNASSPIKYYNVTETVKAVFQYGYK 123
               :|   ||:  |:  :| | :    ||      :          | |  :||  ::   ||
Db         58 RETD--GSEHICVSAKVSENQPNGSDYEFVQEYRLGT-----KEQNTRKTV-TLYATPYK 109

Qy        124 NIRNAIEYQVGGGLNITD----------TLIFTDGELCDVFYV--PNADQGCELWV-KKS 170
               :| : |    : ::              ||::|   ||:  |   |:  |||::   |:
Db        110 TEMHATQRQNNNAMRVSQKKDAQDGKGYQLIYSDYNKCDILRVLRENSGHDCELYLHSKA 169

Qy        171 HYKHVPDYCTFVFNVFCAKDRKTYDIFNEECVYNGEPW 208
                  ||  |  |: : | ||  :|    :: ||   ||
Db        170 LDDGVPRECESVYGIACGKDEPSY----KQRVY--YPW 201


RESULT 3
Q8MVC3
ID   Q8MVC3      PRELIMINARY;      PRT;   306 AA.
AC   Q8MVC3;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Putative secreted histamine binding protein.
OS   Ixodes scapularis (Black-legged tick) (Deer tick).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC   Parasitiformes; Ixodida; Ixodidae; Ixodes.
OX   NCBI_TaxID=6945;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Rhode Island; TISSUE=Salivary gland;
RA   Valenzuela J.G., Francischetti I.M., Pham V.M., Garfield M.,
RA   Mather T.N., Ribeiro J.M.C.;
RT   "Exploring the Sialome of the Tick Vector of Lyme Disease, Ixodes
RT   scapularis.";
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF483717; AAM93639.1; -.
SQ   SEQUENCE   306 AA;  34232 MW;  B2DB09181CC4395D CRC64;

  Query Match             10.4%;  Score 117.5;  DB 5;  Length 306;
  Best Local Similarity   20.0%;  Pred. No. 0.027;
  Matches   46;  Conservative   34;  Mismatches   73;  Indels   77;  Gaps    7;

Qy          1 MKMQVVLLLTFVSAALATQAETTSAKAGENPLWAHEELLGKYQDAWKSIDQGVSVTYVLA 60
              | :    |   :::| | ||: |      |||  :|  ||  | |||:|::     :||
Db          1 MGLMYAALFACIASASAWQAQMTIRNPENNPL-LNEPSLGPLQSAWKAINKSKDDPFVLM 59

Qy         61 ----------------------------KTTYENDTGSWGSQFKCLQVQEIERKEEDYT 91
                                          || |:|     :   |: |     : ||
Db         60 FRSRNHEPNITCVVVTASLHNETLKIVNFTRTYYNET---DGKNDTLEYQVRALNQTDYK 116

Qy         92 VTSV-------------------------FTFRNASSPIKYYNVTETVKAVFQYGYKNI 125
              : :|                       ::   :| |:     |: :||
Db        117 LENVIRAGLKGTPSDKPTPLGSNMYIEYGDYSCNTSSKPL-----TDMLKAA-------- 163

Qy        126 RNAIEYQVGG-----GLNITDTLIFTDGELCDVFYVPNADQGCELWVKKS 170
              ::|:  : |      |:|  |  :  :   |::   |    ||: |::||
Db        164 KDAVGSEAGSAEPVEGVNYLDFYVVHNQPHCNILRSPLLKGGCDFWLRKS 213


RESULT 4
Q8I9T9
ID   Q8I9T9      PRELIMINARY;      PRT;   176 AA.
AC   Q8I9T9;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   TSGP4.
OS   Ornithodoros savignyi.
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC   Parasitiformes; Ixodida; Argasidae; Ornithodoros.
OX   NCBI_TaxID=69826;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Salivary gland;
RX   MEDLINE=21317941; PubMed=11425229;
RA   Mans B.J., Venter J.D., Vrey P.J., Louw A.I., Neitz A.W.;
RT   "Identification of putative proteins involved in granule biogenesis of
RT   tick salivary glands.";
RL   Electrophoresis 22:1739-1746(2001).
DR   EMBL; AF452891; AAN76831.1; -.
SQ   SEQUENCE   176 AA;  19375 MW;  44B7559F3BCB6A25 CRC64;

  Query Match             9.9%;  Score 111.5;  DB 5;  Length 176;
  Best Local Similarity   19.2%;  Pred. No. 0.044;
  Matches   30;  Conservative   33;  Mismatches   64;  Indels   29;  Gaps    6;

Qy         44 DAWKSIDQGVSVTYVLAKTTYENDTGSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASS 103
              | | :: :|    ::: | |||           ||: ::     |  :|:  :  :  | :
Db         23 DVW-NVLKGSDSKFLMVKRTYERGAN------KCVYMKRTSMDESSHTLEVLMGYSKAGT 75

Qy        104 PIKY-----YNVTETVKAVFQYGYKNIR-----NAIEYQVGGGLNITDTLIFTDGELCDV 153
                :      | || | :    |    :|     : ::::          |:::| : |::
Db         76 TTDFVEPSKYTVTATSEGASTYNMMTVRRGPASHGVKFE----------LVYSDDQGCNI 125

Qy        154 FYVFNA--DQGCELWVKKSHYKHVPDYCTFVFNVFC 187
                :  :     ||||  :   |:|   |:  |   |
Db        126 LQMKTSPFPGKCELWAPEGKAKNVESSCSGKFKELC 161


RESULT 5
Q8MVC2
ID   Q8MVC2      PRELIMINARY;      PRT;   311 AA.
AC   Q8MVC2;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Putative secreted histamine binding protein.
OS   Ixodes scapularis (Black-legged tick) (Deer tick).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC   Parasitiformes; Ixodida; Ixodidae; Ixodes.
```

```
OX   NCBI_TaxID=6945;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Rhode Island; TISSUE=Salivary gland;
RA   Valenzuela J.G., Francischetti I.M., Pham V.M., Garfield M.,
RA   Mather T.N., Ribeiro J.M.C.;
RT   "Exploring the Sialome of the Tick Vector of Lyme Disease, Ixodes
RT   scapularis.";
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF483718; AAM93640.1; -.
SQ   SEQUENCE   311 AA;  36581 MW;  D7E529FEFE4CFA9A CRC64;

  Query Match             9.2%;  Score 103;  DB 5;  Length 311;
  Best Local Similarity   19.6%;  Pred. No. 0.54;
  Matches   59;  Conservative   35;  Mismatches   79;  Indels  128;  Gaps   12;

Qy         12 VSAALATQAETTSAKAGENPLWAHEELLGKYQDAWKSIDQGVSVTYVLAKTTYENDTGSW 71
              |||| |      :  :| :|    : : ||  ||||::|    : :| |  ::       |
Db         13 VSAAYAEVEFQSWDRAPDNNPDLNRKDLGAMQDAWRTIKFTANHSYYLIYSS------GW 66

Qy         72 GSQ-----FKCLQVQEIERKEEDYTVTSVFTFRNASSPIKYYN--------VTETVKAVF 118
              |::      :||||   :           :| ::|:   |:||          |: |:|
Db         67 GTREHYEDVRCLQVHSSDLN---------YTLKSANYTSKWYNRTSKRMNSSTQYVQAAK 117

Qy        119 Q--YGYKNIR------------NAIEYQV-------GGGLNI-------------TDTLI 144
              |  |  :||              |   | :         || |               ::  :
Db        118 QKYYSIENIMHLGQPQREVTSPNGTCYNLNFNFLCESGGCRIHHQECWQKRWTKYSEKYV 177

Qy        145 FTDGELCDVFYVPNADQG---CELW--------------------------VKKSH---- 171
                 || |      |:|   || |                           ::||
Db        178 LFSTPLCYVVRSLQDDEGYESCEFWLSEDWLKKNVTIPQVVTILEKEDSDEIEKSEEREE 237

Qy        172 --------------------------YKHVPDYCTFVFNVFCA------KDRKTYDIFN 198
                                       :| :|  | : | : |         |:|  |  |
Db        238 SYERKESVQCEEEPRTANKTTFLYDLLFKELPSSCRYAFLLNCGYPKYRIYDKKDCDKIN 297

Qy        199 E 199
              |
Db        298 E 298


RESULT 6
Q837S0
ID   Q837S0      PRELIMINARY;      PRT;   722 AA.
AC   Q837S0;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Amino acid ABC transporter, amino acid-binding/permease protein.
GN   EF0761.
OS   Enterococcus faecalis (Streptococcus faecalis).
OC   Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX   NCBI_TaxID=1351;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=V583 / ATCC 700802;
RX   MEDLINE=22550857; PubMed=12663927;
RA   Paulsen I.T., Banerjei L., Myers G.S.A., Nelson K.E., Seshadri R.,
RA   Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
RA   Tettelin H., Dodson R.J., Umayam L., Brinkac L., Beanan M.,
RA   Daugherty S., DeBoy R.T., Durkin S., Kolonay J., Madupu R., Nelson W.,
RA   Vamathevan J., Tran B., Upton J., Hansen T., Shetty J., Khouri H.,
RA   Utterback T., Radune D., Ketchum K.A., Dougherty B.A., Fraser C.M.;
RT   "Role of mobile DNA in the evolution of vancomycin-resistant
RT   Enterococcus faecalis.";
RL   Science 299:2071-2074(2003).
DR   EMBL; AE016949; AAO80578.1; -.
DR   TIGR; EF0761; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. . .; IEA.
DR   GO; GO:0005234; F:glutamate-gated ion channel activity; IEA.
DR   GO; GO:0004970; F:inotropic glutamate receptor activity; IEA.
DR   GO; GO:0005215; F:transporter activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR000515; BPD_transp.
DR   InterPro; IPR001320; Ion_glu_receptor.
DR   InterPro; IPR001311; SBP/glu_receptor.
DR   InterPro; IPR001638; SBP_bac_3.
DR   Pfam; PF00528; BPD_transp; 1.
DR   Pfam; PF00497; SBP_bac_3; 2.
DR   SMART; SM00062; PBPb; 2.
DR   SMART; SM00079; PBPe; 1.
DR   PROSITE; PS00402; BPD_TRANSP_INN_MEMBR; 1.
KW   Complete proteome.
SQ   SEQUENCE   722 AA;  79713 MW;  D7F76B3B20BDFE62 CRC64;

  Query Match             9.1%;  Score 102.5;  DB 16;  Length 722;
  Best Local Similarity   26.9%;  Pred. No. 1.8;
  Matches   60;  Conservative   26;  Mismatches   72;  Indels   65;  Gaps   15;

Qy         16 LATQAETTSA-----KAGENPLWAHEELLGKYQDAWKSI-DQGVSVTYVLAKTTY----- 64
              |     || |:      | |:||     ||: |:    |:: | |   ||      |
Db        204 LVGDKETGSSYGFAVKKGQNP-----ELIKKFNAGLKNLKDNG---TYDKILNNYLATGD 255

Qy         65 ENDTGSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASSPIKYYNV--------TETVKA 116
              | :|   | | |     :|  |:| | : |  ||    :| :: |          : ||
Db        256 ETNTQDAGEQMK-----KITPKKEKYVIASDSTF----APFEFQNAQGDYVGIDVDLVKR 306

Qy        117 V---------FQY-GYKNIRNAIEY-QVGG---GLNITDTLIFTDGELCDVFYVPNADQG 162
                        |:: |: :   |:|  |  |   |: |||      | :    | ||  | |
Db        307 AAELQGFTVEFKFIGFSSAVQAVESGQADGMVAGMTITD-----DRKKAFDFSVPYFDSG 361

Qy        163 CELWVKKSH--YKHVPDYCTFVFNV--------FCAKDRKTYD 195
               :: ||| :    |   |        |        |  |::| ||
Db        362 IQIAVKKGNDKIKSYDDLKGKKVGVKIGTESADFLEKNKKKYD 404


RESULT 7
Q95WZ5
ID   Q95WZ5      PRELIMINARY;      PRT;   221 AA.
AC   Q95WZ5;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Histamine binding protein.
GN   HBP.
OS   Ixodes scapularis (Black-legged tick) (Deer tick).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
OC   Parasitiformes; Ixodida; Ixodidae; Ixodes.
OX   NCBI_TaxID=6945;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Das S., Banerjee G., DePonte K., Marcantonio N., Kantor F.S.,
RA   Fikrig E.;
RT   "Salp25D, an Ixodes scapularis antioxidant, is one of 14
RT   immunodominant antigens in engorged tick salivary glands.";
RL   J. Infect. Dis. 184:0-0(2001).
DR   EMBL; AF209913; AAK97816.1; -.
SQ   SEQUENCE   221 AA;  25359 MW;  7B995AC9C0AFA252 CRC64;

  Query Match             8.7%;  Score 97.5;  DB 5;  Length 221;
  Best Local Similarity   23.4%;  Pred. No. 1.1;
  Matches   49;  Conservative   43;  Mismatches   80;  Indels   37;  Gaps   14;

Qy         10 TFVSAALATQAET-------TSAKAGENPLWAHEELLGKY--QDAWKSIDQGVSVTYVLA 60
              |: |    |:  |       |:|   |:|          ||  |:| : ::  :: | :
Db         31 TYGSTGTTTRPGTRGARMIVTTAPPEEDP--------SKYKEQNATRVVE--MNATQWVK 80

Qy         61 KTTYENDTGSWGSQFKCLQVQEIERKEEDYTVTSV-FTFRNASSPIKYYNVTET--VKAV 117
                ||:  |    |: :|    : :|::    | |:  | :|  |     :  : ||  :| :
Db         81 WRTYD-VTDFSGNPVQCENFRVMEKR----TPTNYSFQYRYKSKN-SWETIDETLILKDI 134

Qy        118 FQYGYKNIRNAIEYQVGGGLNITDTLI-FTDGELCDVFYVPNADQG---CELWVKK-SHY 172
               ::|: |: |     :|     || |: :::   | |  :|  :||   |:||:   :
Db        135 GEHGFPNVMNFQRTPIG---IATDNLVLYSNYVNCTVLRIPFTNQGERHCDLWMANLTLS 191
```

```
Qy        173 KHVPDYCTFVFNVFCAKDRKTYDIFNEEC 201
              :  || |   |  :|   : | ::   |
Db        192 QETPDDCLNKFFEYC-NTTQIYRVYYPSC 219


RESULT 8
Q8A286
ID   Q8A286      PRELIMINARY;      PRT;   508 AA.
AC   Q8A286;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein.
GN   BT3420.
OS   Bacteroides thetaiotaomicron.
OC   Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC   Bacteroidaceae; Bacteroides.
OX   NCBI_TaxID=818;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VPI-5482 / ATCC 29148;
RX   MEDLINE=22550858; PubMed=12663928;
RA   Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
RA   Chiang H.C., Hooper L.V., Gordon J.I.;
RT   "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
RL   Science 299:2074-2076(2003).
DR   EMBL; AE016940; AAO78526.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   508 AA;  59877 MW;  B80E9E38262D6450 CRC64;

  Query Match             8.4%;  Score 95;  DB 16;  Length 508;
  Best Local Similarity   21.9%;  Pred. No. 5.3;
  Matches   43;  Conservative   27;  Mismatches   64;  Indels   62;  Gaps    9;

Qy          6 VLLLTFVSAALATQ-------AETTSAKAGENPLWAHEELLGKYQDAWKSID----QGVS 54
              ::  | |       ||         ||  |||| |            ::  |::|    || |
Db          8 IVSLFFFYTGCDTQQTKYPFSVETVLAKAGNN-----------RKELEKALDYFYKQGDS 56

Qy         55 VT------YVLAKTTYENDTGSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASSPIKYY 108
              :       |     : ::|  | ::          | |  |:             |   |
Db         57 IKIKAIEFLVAHMDIHYSETYYWKTR---------EGKRVDF-----------SEFDYA 95

Qy        109 NVTETVKAVFQYGYKNIRNAIEYQVGGGLNITDTLIFTDGELCDVFYVPNADQGCELWVK 168
              |:   |||:      |          | |:  ||:|:     |   : : | :   : | :
Db         96 NLETAVKAIDSMRKKY----------GSLDFQDTIIYDVNSLTGKYLINNVNHAVDTW-R 144

Qy        169 KSHYKHVP--DYCTFV 182
              | || :|  |:| ::
Db        145 LSEYKDIPFNDFCEYI 160


RESULT 9
Q8IDR7
ID   Q8IDR7      PRELIMINARY;      PRT;   972 AA.
AC   Q8IDR7;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein.
GN   PF13_0230.
OS   Plasmodium falciparum (isolate 3D7).
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=36329;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Harris B., Lennard N., Clark L., Line A., Barron A., Corton C.,
RA   Berriman M., Pain A., Hall N., Atkin R., Chillingworth C., Doggett J.,
RA   Ormond D., Sanders M., Hayes R., Hall S., Quail M., Barrell B.;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL844509; CAD52552.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   972 AA;  119242 MW;  70471A60C0327C32 CRC64;

  Query Match             7.9%;  Score 89;  DB 5;  Length 972;
  Best Local Similarity   19.8%;  Pred. No. 43;
  Matches   43;  Conservative   33;  Mismatches   81;  Indels   60;  Gaps    9;

Qy          4 QVVLLLTFVSAALATQAETTSAKAGENPLWAHEELLGKYQDAWKSIDQGV-----SVTYV 58
              :: ||  ::| :: :|     |  |         : | :   ||:|  :        ::
Db        289 KICLLFIYISCSIVSQTYKHIMKKENN--------VEKKRKKMKSLDNFIIKNINECKHI 340

Qy         59 LAKTTYENDTGSWGSQFKCLQVQEIE----RKEEDYTVTSVFTFRNASSPIKYYNVTETV 114
              |                ||:  | |    |::   |  ::   :   : ||| |:
Db        341 LMDNLIN----------KCINDQYIRYIILRRKYCSTYLNINVLKKYLNSIKYLNIIINIK 390

Qy        115 KAVFQ--YGYKN---IRNAIEYQVGGGLNITDTLIFTDGELCD-VFYVPNADQGCELWVK 168
              |: :   | | |       :| |                   |  :||:       |
Db        391 KSFYYICYLYYNSYIFNKSIYY------------------CSLIFYL--------LQKY 423

Qy        169 KSHYKHVPDYCTFVFNVFCAKDRKTYDIFNEECVYNG 205
                || ::       ||:|  |||| |: :|:  :  |
Db        424 NLHYTYLYQILLIKFNLFLLKDRKIYNDYNKHILLYG 460


RESULT 10
Q939L2
ID   Q939L2      PRELIMINARY;      PRT;   598 AA.
AC   Q939L2;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Arylsulfate sulfotransferase.
GN   ASTA.
OS   Citrobacter freundii.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Citrobacter.
OX   NCBI_TaxID=546;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MB4-8242;
RX   MEDLINE=21425731; PubMed=11534764;
RA   Kang J.W., Jeong Y.J., Kwon A.R., Yun H.J., Kim D.H., Choi E.C.;
RT   "Cloning, sequence analysis, and characterization of the astA gene
RT   encoding an arylsulfate sulfotransferase from Citrobacter freundii.";
RL   Arch. Pharm. Res. 24:316-322(2001).
DR   EMBL; AY029222; AAK31641.1; -.
DR   GO; GO:0016740; F:transferase activity; IEA.
KW   Transferase.
SQ   SEQUENCE   598 AA;  66501 MW;  4CE9BB26545F62A5 CRC64;

  Query Match             7.7%;  Score 87;  DB 2;  Length 598;
  Best Local Similarity   24.0%;  Pred. No. 34;
  Matches   46;  Conservative   25;  Mismatches   55;  Indels   66;  Gaps   10;

Qy         27 AGENPLWAHEE-----------LLGKYQDAWKSIDQGVSVTYVLAKTTYEN-------- 66
              || |  |||               |  ::|   | | :   | ::|| :   |
Db        377 AGRN--WAHVNSIAYDAKDDSIILSSRHQGVVK-IGRDKQVKWILAPSKGWNKALASKLL 433

Qy         67 -DTGSWGSQFKCLQVQEIERKEEDYTVT----------SVFTFRN---------ASSPIK 106
                |:  || :  : |  : |:| |              ::  | |         |  :|
Db        434 KPVDDKGNALKCDENGKCENTDFDFTYTQHTAWLSSKGTLTIFDNGDGRGLEQPALPTMK 493

Qy        107 Y-----YNVTE---TVKAVFQYG-------YKNIRNAIEYQ--------VGGGLNITDTL 143
              |     | : |   ||: |::||       |  | : |||          || :|: |
Db        494 YSRFVEYKIDEKKGTVQQVWEYGKERGYDFYSPITSVIEYQKDRDTMFGFGGSINLFDVG 553

Qy        144 IFTDGELCDVFY 155
                | |:: :: |
Db        554 QPTIGKINEIDY 565


RESULT 11
Q92N47
ID   Q92N47      PRELIMINARY;      PRT;   932 AA.
AC   Q92N47;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
```

```
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Putative outer membrane receptor protein.
GN   R02383 OR SMC02721.
OS   Rhizobium meliloti (Sinorhizobium meliloti).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX   NCBI_TaxID=382;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=1021;
RX   MEDLINE=21396507; PubMed=11481430;
RA   Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA   Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA   Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA   Pohl T., Portetelle D., Puehler A., Purnelle B., Ramsperger U.,
RA   Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RT   "Analysis of the chromosome sequence of the legume symbiont
RT   Sinorhizobium meliloti strain 1021.";
RL   Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
DR   EMBL; AL591790; CAC46962.1; -.
DR   GO; GO:0016020; C:membrane; IEA.
DR   GO; GO:0004872; F:receptor activity; IEA.
DR   GO; GO:0005215; F:transporter activity; IEA.
DR   GO; GO:0006810; P:transport; IEA.
DR   InterPro; IPR000531; TonB_boxC.
DR   Pfam; PF00593; TonB_dep_Rec; 1.
KW   Complete proteome.
SQ   SEQUENCE   932 AA;  101872 MW;  E4F6473BAF5AC484 CRC64;

  Query Match             7.7%;  Score 86.5;  DB 16;  Length 932;
  Best Local Similarity   25.5%;  Pred. No. 68;
  Matches   41;  Conservative   24;  Mismatches   47;  Indels   49;  Gaps   11;

Qy         44 DAWKSIDQGVSVTYVLAKTTYENDTG----SWGSQ--FKCLQ-VQEIERKEEDYTVTSVF 96
              | :| ||:|               |||      || :  |:      |  | | ::|  ||
Db        221 DGFKDIDRG-------------GDTGFDLTDWGGKLGFRTADGTQNFEFKYQNYDETSNE 267

Qy         97 TFRNAS------SPIKYYN-----VTETVKAVFQYGYKNIRNAIEYQVGGGLNITDTLIF 145
              |:   :      :| : ||      | :    ::|: :         |::   :|:| |:
Db        268 TYLGLTLDDFKDTPFRRYNASQKDVIDAAHNIYQFSH-------SYEITPDINLTTTVYR 320

Qy        146 TDG-----ELCDVFYVPNADQGCELWVKKSHYKHVPD-YCT 180
              ||      :| ||    ::| |   ||  |:    || | |
Db        321 TDTTRAWYKLNDV--RNDSDTG---WVSLSNILADPDTYST 356


RESULT 12
Q8XMD0
ID   Q8XMD0      PRELIMINARY;      PRT;   399 AA.
AC   Q8XMD0;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein CPE0759.
GN   CPE0759.
OS   Clostridium perfringens.
OC   Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC   Clostridium.
OX   NCBI_TaxID=1502;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=13 / Type A;
RX   MEDLINE=21664373; PubMed=11792842;
RA   Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RA   Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT   "Complete genome sequence of Clostridium perfringens, an anaerobic
RT   flesh-eater.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
DR   EMBL; AP003188; BAB80465.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   399 AA;  46802 MW;  61019F007582B815 CRC64;

  Query Match             7.6%;  Score 86;  DB 16;  Length 399;
  Best Local Similarity   23.8%;  Pred. No. 24;
  Matches   30;  Conservative   21;  Mismatches   37;  Indels   38;  Gaps    4;

Qy         42 YQDAWKS---IDQG------------VSVTYVLAKTTYENDTGSWGSQFKCLQVQEIER 85
              |::| ||   |: |             :::::   : | |   |:   || | : ::
Db        224 YEEAKKSLEYIESGFNLMCSRYSENIEMNLSFSRIRETNSNIKASYKEAFKALSMIDM-- 281

Qy         86 KEEDYTVTSVFTFRNASSPIKYYNV--------TETVKAVFQYGYKNIRNAIEYQVGGGL 137
                        : | |  ||| :           :| | || |:||   :||    |:
Db        282 -----------YDNDSKIIKYNELGIIRLFIELAEIKDVHQYCYENIGPILEYDKSHGM 329

Qy        138 NITDTL 143
              |: :||
Db        330 NLLETL 335


RESULT 13
Q51717
ID   Q51717      PRELIMINARY;      PRT;   584 AA.
AC   Q51717;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein precursor.
OS   Pseudomonas fluorescens.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=294;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DSM 7139;
RX   MEDLINE=95050280; PubMed=7961472;
RA   Schirmer A., Jendrossek D.;
RT   "Molecular characterization of the extracellular poly(3-
RT   hydroxyoctanoic acid) [P(3HO)] depolymerase gene of Pseudomonas
RT   fluorescens GK13 and of its gene product.";
RL   J. Bacteriol. 176:7065-7073(1994).
DR   EMBL; U10470; AAA64539.1; -.
KW   Hypothetical protein; Signal.
FT   SIGNAL        1      26       POTENTIAL.
SQ   SEQUENCE   584 AA;  63249 MW;  3AA16FAE898D2CBF CRC64;

  Query Match             7.6%;  Score 86;  DB 2;  Length 584;
  Best Local Similarity   27.1%;  Pred. No. 40;
  Matches   29;  Conservative   15;  Mismatches   51;  Indels   12;  Gaps    2;

Qy         39 LGKYQDAWKSIDQGVSVTYVLAKTTYENDTGSWGSQFKCL-------QVQEIERKEEDYT 91
              :| : | ::  |       |  :     | ||  |:     |       :|| :|   : |
Db        433 VGTHLDGYREKDM-----YQASLAAIYNFTGVLGADSATLFTELVASRVQGLESDVDYYE 487

Qy         92 VTSVFTFRNASSPIKYYNVTETVKAVFQYGYKNIRNAIEYQVGGGLN 138
               ||  |   ||  : | ||   |  |  :||:   | :  |: ||:
Db        488 ATSTATGAQASLQLTYTNVLNLVNLVPSFGYQYSINGVAPQLTNGLD 534


RESULT 14
O33983
ID   O33983      PRELIMINARY;      PRT;   917 AA.
AC   O33983;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   AMI.
GN   AMI.
OS   Listeria monocytogenes.
OC   Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX   NCBI_TaxID=1639;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=EGD;
RX   MEDLINE=97426036; PubMed=9282740;
RA   Braun L., Dramsi S., Dehoux P., Bierne H., Lindhal G., Cossart P.;
```

```
RT   "InlB : an invasion protein of Listeria monocytogenes with a novel
RT   type of surface association.";
RL   Mol. Microbiol. 25:285-294(1997).
DR   EMBL; U82488; AAC45605.1; -.
DR   GO; GO:0008745; F:N-acetylmuramoyl-L-alanine amidase activity; IEA.
DR   GO; GO:0009253; P:peptidoglycan catabolism; IEA.
DR   InterPro; IPR002502; Amidase_2.
DR   Pfam; PF01510; Amidase_2; 1.
DR   SMART; SM00644; Ami_2; 1.
SQ   SEQUENCE   917 AA;  102352 MW;  21EA2452B4F103B7 CRC64;

  Query Match             7.6%;  Score 85.5;  DB 2;  Length 917;
  Best Local Similarity   23.2%;  Pred. No. 81;
  Matches   38;  Conservative   24;  Mismatches   49;  Indels   53;  Gaps    9;

Qy         28 GENPLWAHEELLGKYQDAWKSIDQGVSVTYVLAKTTYENDT---GSWG-----------S 73
              |:  :|:|:                :||  |  ||: :       ||
Db        223 GKGTVWSHD----------------AVTRYLGGTTHTDPVAYFNQWGYNFNNFVSLINE 265

Qy         74 QFKCLQVQEIERKEEDYTVTSVFTFRNASSPIKYY--NVTETVKAV---FQYGYKNIR-- 126
              ::| :||   |: | |  :|:   : |:    :   | ||  | |     |  ||:|
Db        266 KYKAMQV-NYEKIEYDKAITAYSRVKTATGNSVWTKPNKTEGAKLVNPLSSYSGKNLRII 324

Qy        127 -------NAIEYQ--VGGGLNITDTLIFTDGELCDVFYVPNADQ 161
                    | ||  |||       |: : | :  : || |: ::
Db        325 REAKTSGGTIWYQFSVGG-----KTIGWVDSKALNTFYTPSMEK 363


RESULT 15
Q8A8I2
ID   Q8A8I2      PRELIMINARY;      PRT;  1090 AA.
AC   Q8A8I2;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   OmpA-related protein.
GN   BT1185.
OS   Bacteroides thetaiotaomicron.
OC   Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC   Bacteroidaceae; Bacteroides.
OX   NCBI_TaxID=818;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VPI-5482 / ATCC 29148;
RX   MEDLINE=22550858; PubMed=12663928;
RA   Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
RA   Chiang H.C., Hooper L.V., Gordon J.I.;
RT   "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
RL   Science 299:2074-2076(2003).
DR   EMBL; AE016930; AAO76292.1; -.
DR   GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR   GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR   InterPro; IPR008969; CarboxypepD_reg.
DR   InterPro; IPR001064; Crystallin.
DR   InterPro; IPR000111; Glyco_hydro_GHD.
DR   ProDom; PD002572; Glyco_hydro_GHD; 1.
DR   PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
KW   Complete proteome.
SQ   SEQUENCE   1090 AA;  122044 MW;  63C8B3A1E8A2DFA6 CRC64;

  Query Match             7.6%;  Score 85.5;  DB 16;  Length 1090;
  Best Local Similarity   23.8%;  Pred. No. 1e+02;
  Matches   44;  Conservative   19;  Mismatches   81;  Indels   41;  Gaps    9;

Qy         47 KSIDQG----VSVTYVLAKTTYEN----DTGSWGSQFKCLQVQEIERKEEDYTVTSV--F 96
              |  | |    :  || :||    |       :| :    : |  :      |:: |
Db        821 KKYDFGFTGSIGYTYTMAKDLTANPGSAPNSAWQNN---VAVNSLNDPGVSYSLFSTPHR 877

Qy         97 TFRNASSPIKYYNVTETVKAVFQYGYKNIRNAIEY--QVGGGLNITDTLIFTDGELCDVF 154
                |||  | |    :|  ::|  ||   | :  |   : |                ||   |:
Db        878 IIANASYEINYAKCLKTTFSLFYSGYHTGRYSYTYYNDMNG-----------DGNYSDLI 926

Qy        155 YVPNADQGCEL--------WVKKSHYKHVPDYCTFVFNVFCAKDRKTYDIFNEECVYNGE 206
                |||| :               :  |      |:  || |    ||||      :    ||
Db        927 YVPNSQDEMTFVDITDKSGAITYSAVDQAKDFWDFVNNDSYLKDRK-----GKYVERNGS 981

Qy        207 --PWL 209
                ||:
Db        982 LTPWI 986


Search completed: March  9, 2004, 13:35:31
Job time : 50 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: March 9, 2004, 13:30:30 ; Search time 60 Seconds
(without alignments)
984.207 Million cell updates/sec

Title: US-09-555-296B-4
Perfect score: 1125
Sequence: 1 MKMQVVLLLTFVSAALATQA.....DRKTYDIFNEECVYNGEPWL 209

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1125	100.0	209	2	AAW37449	Aaw37449 Tick vaso
2	1125	100.0	209	2	AAY18081	Aay18081 Histamine
3	364	32.4	203	2	AAY18086	Aay18086 Histamine
4	353	31.4	203	2	AAY18085	Aay18085 Histamine
5	290.5	25.8	190	2	AAW37447	Aaw37447 Tick vaso
6	290.5	25.8	190	2	AAY18079	Aay18079 Histamine
7	284.5	25.3	198	2	AAY18084	Aay18084 Histamine
8	282	25.1	171	4	AAB73261	Aab73261 Histacali
9	282	25.1	171	4	AAB74289	Aab74289 Histacali
10	265	23.6	200	2	AAW37448	Aaw37448 Tick vaso
11	265	23.6	200	2	AAY18080	Aay18080 Histamine
12	252	22.4	190	2	AAW37446	Aaw37446 Tick vaso
13	252	22.4	190	2	AAY18078	Aay18078 Histamine
14	251.5	22.4	182	4	AAB73262	Aab73262 Histacali
15	251.5	22.4	182	4	AAB74290	Aab74290 Histacali
16	239	21.2	172	4	AAB73260	Aab73260 Histacali
17	239	21.2	172	4	AAB74288	Aab74288 Histacali
18	140	12.4	284	2	AAY18088	Aay18088 Histamine
19	127.5	11.3	285	2	AAY18087	Aay18087 Histamine
20	126.5	11.2	321	2	AAY18089	Aay18089 Histamine
21	102.5	9.1	722	4	AAU35203	Aau35203 Enterococ
22	97.5	8.7	221	4	AAY97764	Aay97764 I. scapul
23	96	8.5	650	6	ABU19265	Abu19265 Protein e
24	91	8.1	202	2	AAR37467	Aar37467 Sequence
25	88	7.8	644	6	ABU33830	Abu33830 Protein e
26	87	7.7	639	6	ABU48728	Abu48728 Protein e
27	86	7.6	2778	4	ABB58683	Abb58683 Drosophil
28	85	7.6	254	3	AAG17991	Aag17991 Arabidops
29	85	7.6	411	4	ABG03896	Abg03896 Novel hum
30	85	7.6	437	3	AAG17990	Aag17990 Arabidops
31	85	7.6	441	3	AAG17989	Aag17989 Arabidops
32	85	7.6	656	6	ABU35973	Abu35973 Protein e
33	84.5	7.5	485	6	ABU29629	Abu29629 Protein e
34	84.5	7.5	497	7	ADC95652	Adc95652 E. faeciu
35	84.5	7.5	917	5	ABB48437	Abb48437 Listeria
36	84	7.5	217	4	AAY97773	Aay97773 I. scapul
37	83	7.4	207	2	AAY18083	Aay18083 Histamine
38	83	7.4	2285	2	AAW98149	Aaw98149 Bacillus
39	82	7.3	555	4	ABB66457	Abb66457 Drosophil
40	82	7.3	876	4	ABB66456	Abb66456 Drosophil
41	81.5	7.2	775	6	ABU20485	Abu20485 Protein e
42	81	7.2	659	2	AAW24123	Aaw24123 Protease.
43	81	7.2	659	2	AAW24121	Aaw24121 Thermococ
44	81	7.2	659	2	AAW94840	Aaw94840 WO9856926
45	80.5	7.2	597	2	AAW20536	Aaw20536 H. pylori

ALIGNMENTS

RESULT 1
AAW37449
ID AAW37449 standard; protein; 209 AA.
XX
AC AAW37449;
XX
DT 27-AUG-2003 (revised)
DT 08-JUN-1998 (first entry)
XX
DE Tick vasoactive amine binding protein D.RET6.
XX
KW Vasoactive amine binding protein; D.RET6; histamine; serotonin; assay;
KW antihistamine; anti-inflammatory; insect bite; snake bite; scorpion bite;
KW dermatitis; vaccine; transgenic animal; tick.
XX
OS Dermacentor reticulatus.
XX
FH Key Location/Qualifiers
FT Peptide 1. .28
FT /label= Sig_peptide
XX
PN WO9744451-A2.
XX
PD 27-NOV-1997.
XX
PF 19-MAY-1997; 97WO-GB001372.
XX
PR 18-MAY-1996; 96GB-00010484.
PR 18-APR-1997; 97GB-00007844.
XX
PA (OXFO-) OXFORD VACS LTD.
XX
PI Paesen GC, Nuttall PA;
XX
DR WPI; 1998-018506/02.
DR N-PSDB; AAV00230.
XX
PT New vasoactive amine binding proteins and related nucleic acid, vectors -
PT transformed cells and transgenic animals, used for assaying or removing
PT histamine and as antihistamine or anti-inflammatory agents.
XX
PS Example 2; Fig 4; 44pp; English.
XX
CC This protein comprises tick Dermacenter reticularis (Dr) novel vasoactive
CC amine binding protein (VABP) D.RET6. Its amino acid sequence was deduced
CC from a cDNA clone (see AAV00230) obtained from a Dr salivary gland cDNA
CC library. 3 Novel VASPs, designated FS-HBP1, FS-HBP2 and MS-HBP1 (see
CC AAW37446-48), of the tick Rhipicephalus appendiculatus have also been

CC identified. The VABPs can be expressed in host cells using e.g. a
CC baculovirus expression system. They can be used: (i) to assay histamine
CC (or other VA such as serotonin) in body fluids or cell culture
CC supernatants, e.g. to monitor the effect of allergens; (ii) for binding
CC VA, e.g. to remove histamine from blood, food, cell cultures etc.; (iii)
CC as an antihistamine or anti-inflammatory agents, e.g. for treating
CC insect, snake or scorpion bites or dermatitis, or as a carrier for slow
CC release of histamine-related compounds; (iv) in vaccines to protect
CC against metazoan parasites, especially in animals; (v) as reagents for
CC studying inflammation, involvement of VA in ulcer formation or the immune
CC response etc. VABPs provide a more sensitive assay for histamine than low
CC -affinity antibodies currently used. They may also be more effective and
CC safer than conventional antihistamines. (Updated on 27-AUG-2003 to
CC correct OS field.)
XX
SQ Sequence 209 AA;

Query Match 100.0%; Score 1125; DB 2; Length 209;
Best Local Similarity 100.0%; Pred. No. 6.9e-104;
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKMQVVLLLTFVSAALATQAETTSAKAGENPLWAHEELLGKYQDAWKSIDQGVSVTYVLA 60
Db 1 MKMQVVLLLTFVSAALATQAETTSAKAGENPLWAHEELLGKYQDAWKSIDQGVSVTYVLA 60

Qy 61 KTTYENDTGSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASSPIKYYNVTETVKAVFQY 120
Db 61 KTTYENDTGSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASSPIKYYNVTETVKAVFQY 120

Qy 121 GYKNIRNAIEYQVGGGLNITDTLIFTDGELCDVFYVPNADQGCELWVKKSHYKHVPDYCT 180
Db 121 GYKNIRNAIEYQVGGGLNITDTLIFTDGELCDVFYVPNADQGCELWVKKSHYKHVPDYCT 180

Qy 181 FVFNVFCAKDRKTYDIFNEECVYNGEPWL 209
Db 181 FVFNVFCAKDRKTYDIFNEECVYNGEPWL 209

RESULT 2
AAY18081
ID AAY18081 standard; protein; 209 AA.
XX
AC AAY18081;
XX
DT 06-AUG-1999 (first entry)
XX
DE Histamine binding protein D.RET6.
XX
KW Histamine binding protein; serotonin binding compound; inflammation;
KW gastric acid secretion; allergy; type I hypersensitivity reaction;
KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;
KW drug allergy; abnormal blood pressure; psychological disorder; vaccine;
KW respiratory disease; coronary heart disease; cellular growth regulator;
KW tissue repair; blood-sucking ectoparasite; therapy.
XX
OS Rhipicephalus appendiculatus.
XX
PN WO9927104-A1.
XX
PD 03-JUN-1999.
XX
PF 26-NOV-1998; 98WO-GB003530.
XX
PR 26-NOV-1997; 97GB-00025046.
PR 26-JUN-1998; 98GB-00013917.
XX
PA (OXFO-) OXFORD VACS LTD.
XX
PI Nuttall PA, Paesen GC;
XX
DR WPI; 1999-357841/30.
DR N-PSDB; AAX76967.
XX

PT Histamine and serotonin binding compounds useful for the treatment of
PT allergies.
XX
PS Claim 1; Fig 4; 84pp; English.
XX
CC This sequence is an example of a histamine or serotonin binding compound
CC (A), of the invention. The compounds are useful for regulating the action
CC of histamine and serotonin (in e.g. inflammation and gastric acid
CC secretion), the detection, quantification and removal of histamine or
CC serotonin (in animals, plants, cell cultures, food materials, or humans)
CC and in the treatment of various diseases and allergies (e.g. type I
CC hypersensitivity reactions, urticaria, asthma, allergic rhinitis (hay
CC fever), atopic dermatitis, insect bites and food and drug allergies,
CC abnormal blood pressure, migraine, psychological disorders, respiratory
CC disease, and coronary heart disease). Histamine may also be used to
CC regulate cellular growth and tissue repair. The molecules may also be
CC used as components of vaccines directed against blood-sucking
CC ectoparasites
XX
SQ Sequence 209 AA;

Query Match 100.0%; Score 1125; DB 2; Length 209;
Best Local Similarity 100.0%; Pred. No. 6.9e-104;
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKMQVVLLLTFVSAALATQAETTSAKAGENPLWAHEELLGKYQDAWKSIDQGVSVTYVLA 60
Db 1 MKMQVVLLLTFVSAALATQAETTSAKAGENPLWAHEELLGKYQDAWKSIDQGVSVTYVLA 60

Qy 61 KTTYENDTGSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASSPIKYYNVTETVKAVFQY 120
Db 61 KTTYENDTGSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASSPIKYYNVTETVKAVFQY 120

Qy 121 GYKNIRNAIEYQVGGGLNITDTLIFTDGELCDVFYVPNADQGCELWVKKSHYKHVPDYCT 180
Db 121 GYKNIRNAIEYQVGGGLNITDTLIFTDGELCDVFYVPNADQGCELWVKKSHYKHVPDYCT 180

Qy 181 FVFNVFCAKDRKTYDIFNEECVYNGEPWL 209
Db 181 FVFNVFCAKDRKTYDIFNEECVYNGEPWL 209

RESULT 3
AAY18086
ID AAY18086 standard; protein; 203 AA.
XX
AC AAY18086;
XX
DT 06-AUG-1999 (first entry)
XX
DE Histamine binding protein Ih/Bm-HBP2.
XX
KW Histamine binding protein; serotonin binding compound; inflammation;
KW gastric acid secretion; allergy; type I hypersensitivity reaction;
KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;
KW drug allergy; abnormal blood pressure; psychological disorder; vaccine;
KW respiratory disease; coronary heart disease; cellular growth regulator;
KW tissue repair; blood-sucking ectoparasite; therapy.
XX
OS Boophilus microplus.
OS Ixodes hexagonus.
OS Synthetic.
XX
PN WO9927104-A1.
XX
PD 03-JUN-1999.
XX
PF 26-NOV-1998; 98WO-GB003530.
XX
PR 26-NOV-1997; 97GB-00025046.
PR 26-JUN-1998; 98GB-00013917.
XX
PA (OXFO-) OXFORD VACS LTD.

XX
PI Nuttall PA, Paesen GC;
XX
DR WPI; 1999-357841/30.
DR N-PSDB; AAX76971.
XX
PT Histamine and serotonin binding compounds useful for the treatment of
PT allergies.
XX
PS Claim 14; Fig 8; 84pp; English.
XX
CC This sequence is an example of a histamine or serotonin binding compound
CC (A), of the invention. cDNA encoding this sequence was isolated from a
CC mixed Boophilus microplus/Ixodes hexagonus cDNA expression library. The
CC compounds are useful for regulating the action of histamine and serotonin
CC (in e.g. inflammation and gastric acid secretion), the detection,
CC quantification and removal of histamine or serotonin (in animals, plants,
CC cell cultures, food materials, or humans) and in the treatment of various
CC diseases and allergies (e.g. type I hypersensitivity reactions,
CC urticaria, asthma, allergic rhinitis (hay fever), atopic dermatitis,
CC insect bites and food and drug allergies, abnormal blood pressure,
CC migraine, psychological disorders, respiratory disease, and coronary
CC heart disease). Histamine may also be used to regulate cellular growth
CC and tissue repair. The molecules may also be used as components of
CC vaccines directed against blood-sucking ectoparasites
XX
SQ Sequence 203 AA;

Query Match 32.4%; Score 364; DB 2; Length 203;
Best Local Similarity 38.1%; Pred. No. 6.3e-28;
Matches 75; Conservative 38; Mismatches 78; Indels 6; Gaps 5;

```
Qy          6 VLLLTFVSAALATQAETTSAKAGENPLWAHEELLGKYQDAWKSIDQGVSVTYVLAKTTYE 65
              |: || |:||       :|  :  |   | | : ||:||||||||:| |  ||   ::||
Db          8 VVYLTAVTAADQAPPSSTRNEPLEKTTW-HNQTLGRYQDAWKSINQSVGTTYYFLRSTYN 66

Qy         66 NDTGSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASSPIKYYNVTETVKAVFQYGYKNI 125
              ||:  ||  | || |    : |  :||    |::| |   :: :::| | || : || ::
Db         67 NDS-VWGKNFTCLSVTVTSKYESTFTVEYNTTYKNQSQ--QWVSMSENVTAVQEGGY-SV 122

Qy        126 RNAIEYQVGGGLNITDTLIFTDGELCDVFYVPNADQGCELWVKKSHYKHVPDYCTFVFNV 185
             :| |::          ||::|||: ||| |:|  : | ||||:  : :: |  | |:|::
Db        123 KNIIQWTTENNTKFNDTVVFTDGQTCDVLYIPYKEDGYELWVRSEYLQNTPTCCQFIFDL 182

Qy        186 FCAKDRKTYDIFNEECV 202
              |  | ||:|    ||
Db        183 -VALGRTTYNISTPNCV 198
```

RESULT 4
AAY18085
ID AAY18085 standard; protein; 203 AA.
XX
AC AAY18085;
XX
DT 06-AUG-1999 (first entry)
XX
DE Histamine binding protein Ih/Bm-HBP1.
XX
KW Histamine binding protein; serotonin binding compound; inflammation;
KW gastric acid secretion; allergy; type I hypersensitivity reaction;
KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;
KW drug allergy; abnormal blood pressure; psychological disorder; vaccine;
KW respiratory disease; coronary heart disease; cellular growth regulator;
KW tissue repair; blood-sucking ectoparasite; therapy.
XX
OS Boophilus microplus.
OS Ixodes hexagonus.
OS Synthetic.
XX
PN WO9927104-A1.
XX
PD 03-JUN-1999.
XX
PF 26-NOV-1998; 98WO-GB003530.
XX
PR 26-NOV-1997; 97GB-00025046.
PR 26-JUN-1998; 98GB-00013917.
XX
PA (OXFO-) OXFORD VACS LTD.
XX
PI Nuttall PA, Paesen GC;
XX
DR WPI; 1999-357841/30.
DR N-PSDB; AAX76970.
XX
PT Histamine and serotonin binding compounds useful for the treatment of
PT allergies.
XX
PS Claim 13; Fig 7; 84pp; English.
XX
CC This sequence is an example of a histamine or serotonin binding compound
CC (A), of the invention. cDNA encoding this sequence was isolated from a
CC mixed Boophilus microplus/Ixodes hexagonus cDNA expression library. The
CC compounds are useful for regulating the action of histamine and serotonin
CC (in e.g. inflammation and gastric acid secretion), the detection,
CC quantification and removal of histamine or serotonin (in animals, plants,
CC cell cultures, food materials, or humans) and in the treatment of various
CC diseases and allergies (e.g. type I hypersensitivity reactions,
CC urticaria, asthma, allergic rhinitis (hay fever), atopic dermatitis,
CC insect bites and food and drug allergies, abnormal blood pressure,
CC migraine, psychological disorders, respiratory disease, and coronary
CC heart disease). Histamine may also be used to regulate cellular growth
CC and tissue repair. The molecules may also be used as components of
CC vaccines directed against blood-sucking ectoparasites
XX
SQ Sequence 203 AA;

Query Match 31.4%; Score 353; DB 2; Length 203;
Best Local Similarity 36.3%; Pred. No. 7.8e-27;
Matches 74; Conservative 41; Mismatches 81; Indels 8; Gaps 6;

```
Qy          1 MKMQVVLLLTFVSAALATQAETTSAKAG--ENPLWAHEELLGKYQDAWKSIDQGVSVTYV 58
              ||  :: :    :   | ||  :| :    :|  | | : |  ||||||||:| || ||
Db          1 MKALLIAVGYLAAVTAAPQASPSSPRNEPLKNTTW-HSKELKNYQDAWKSINQNVSTTYY 59

Qy         59 LAKTTYENDTGSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASSPIKYYNVTETVKAVF 118
                ::|| ||:  ||  | || |    : |  :||    |::| |   :: ::|| | ||
Db         60 FLRSTYNNDS-VWGKNFTCLSVTVTSKHESTFTVEYNTTYKNQSQ--QWVSMTENVTAVQ 116

Qy        119 QYGYKNIRNAIEYQVGGGLNITDTLIFTDGELCDVFYVPNADQGCELWVKKSHYKHVPDY 178
              : || :::| |::          ||::||||: ||: |:|  : | ||||:  : :: |
Db        117 EEGY-DVKNIIQWTTENNTKFNDTVVFTDGQTCDLLYIPYKENGYELWVRSDYLQNTPTC 175

Qy        179 CTFVFNVFCAKDRKTYDIFNEECV 202
              | |:|::  |  | ||:|   :||
Db        176 CQFIFDL-VALGRTTYNISTPDCV 198
```

RESULT 5
AAW37447
ID AAW37447 standard; protein; 190 AA.
XX
AC AAW37447;
XX
DT 08-JUN-1998 (first entry)
XX
DE Tick vasoactive amine binding protein 2 FS-HBP2.
XX
KW Female-specific vasoactive amine binding protein 1; FS-HCP1; histamine;
KW serotonin; assay; antihistamine; anti-inflammatory; insect bite;
KW snake bite; scorpion bite; dermatitis; vaccine; transgenic animal; tick.
XX
OS Rhipicephalus appendiculatus.
XX
FH Key Location/Qualifiers

```
FT   Peptide          1. .19
FT                    /label= Sig_peptide
XX
PN   WO9744451-A2.
XX
PD   27-NOV-1997.
XX
PF   19-MAY-1997;   97WO-GB001372.
XX
PR   18-MAY-1996;   96GB-00010484.
PR   18-APR-1997;   97GB-00007844.
XX
PA   (OXFO-) OXFORD VACS LTD.
XX
PI   Paesen GC,  Nuttall PA;
XX
DR   WPI; 1998-018506/02.
DR   N-PSDB; AAV00228.
XX
PT   New vasoactive amine binding proteins and related nucleic acid, vectors -
PT   transformed cells and transgenic animals, used for assaying or removing
PT   histamine and as antihistamine or anti-inflammatory agents.
XX
PS   Example 2; Fig 2; 44pp; English.
XX
CC   This protein comprises tick Rhipicephalus appendiculatus (Ra) novel
CC   female-specific histamine binding protein 2 (FS-HBP2). Its amino acid
CC   sequence was deduced from a cDNA clone (see AAV00228) obtained from a
CC   salivary gland cDNA library. FS-HSP1 (see AAW37446) and male-specific
CC   HSP1 (see AAW37448) and a related protein, D.RET6 (see AAW37449) from
CC   Dermacenter reticularis, were also identified. These novel vasoactive
CC   amine binding proteins (VABPs) can be expressed in host cells using e.g.
CC   a baculovirus expression system. They can be used: (i) to assay histamine
CC   (or other VA such as serotonin) in body fluids or cell culture
CC   supernatants, e.g. to monitor the effect of allergens; (ii) for binding
CC   VA, e.g. to remove histamine from blood, food, cell cultures etc.; (iii)
CC   as an antihistamine or anti-inflammatory agents, e.g. for treating
CC   insect, snake or scorpion bites or dermatitis, or as a carrier for slow
CC   release of histamine-related compounds; (iv) in vaccines to protect
CC   against metazoan parasites, especially in animals; (v) as reagents for
CC   studying inflammation, involvement of VA in ulcer formation or the immune
CC   response etc. VABPs provide a more sensitive assay for histamine than low
CC   -affinity antibodies currently used. They may also be more effective and
CC   safer than conventional antihistamines
XX
SQ   Sequence 190 AA;

  Query Match             25.8%;  Score 290.5;  DB 2;  Length 190;
  Best Local Similarity   36.0%;  Pred. No. 1.2e-20;
  Matches   72;  Conservative   32;  Mismatches   79;  Indels   17;  Gaps    9;

Qy          7 LLLTFVSAALATQAETTSAKAGENPLWAHEELLGKYQDAWKSIDQGVSVTYVLAKTTYEN 66
              ||:  ::  ||       |   |  | || |   | :||||||:   |   | : | ||:|
Db          3 LLILSLALVLA-----LSQVKGNQPDWADEAANGAHQDAWKSLKADVENVYYMVKATYKN 57

Qy         67 DTGSWGSQFKCLQVQEIERKEEDYTVTSVFTF-RNASSPIKYYNVTETVKAVFQYGYKNI 125
              |   ||: | |: |   :  |:: :: : | |  || : :::   || | ||  ||| |
Db         58 DP-VWGNDFTCVGVMANDVNEDEKSIQAEFLFMNNADTNMQF--ATEKVTAVKMYGY-NR 113

Qy        126 RNAIEYQVGGGLNITDTLIFTDGELCDVFYVPNAD---QGCELWVKKSHYKHVPDYCTFV 182
               ||  |:  |   || : ::| : ||| |||  |   :| |||   : | ::|  |
Db        114 ENAFRYETEDGQVFTDVIAYSD-DNCDVIYVPGTDGNEEGYELWT--TDYDNIPANCLNK 170

Qy        183 FNVFCAKDRKTYDIFNEECV 202
              || : |  |:| |:|   |:
Db        171 FNEY-AVGRETRDVFTSACL 189


RESULT 6
AAY18079
ID   AAY18079 standard; protein; 190 AA.
XX
AC   AAY18079;
XX
DT   06-AUG-1999  (first entry)
XX
DE   Histamine binding protein FS-HBP2.
XX
KW   Histamine binding protein; serotonin binding compound; inflammation;
KW   gastric acid secretion; allergy; type I hypersensitivity reaction;
KW   asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;
KW   drug allergy; abnormal blood pressure; psychological disorder; vaccine;
KW   respiratory disease; coronary heart disease; cellular growth regulator;
KW   tissue repair; blood-sucking ectoparasite; therapy.
XX
OS   Rhipicephalus appendiculatus.
XX
PN   WO9927104-A1.
XX
PD   03-JUN-1999.
XX
PF   26-NOV-1998;   98WO-GB003530.
XX
PR   26-NOV-1997;   97GB-00025046.
PR   26-JUN-1998;   98GB-00013917.
XX
PA   (OXFO-) OXFORD VACS LTD.
XX
PI   Nuttall PA,  Paesen GC;
XX
DR   WPI; 1999-357841/30.
DR   N-PSDB; AAX76965.
XX
PT   Histamine and serotonin binding compounds useful for the treatment of
PT   allergies.
XX
PS   Claim 1; Fig 2; 84pp; English.
XX
CC   This sequence is an example of a histamine or serotonin binding compound
CC   (A), of the invention. The compounds are useful for regulating the action
CC   of histamine and serotonin (in e.g. inflammation and gastric acid
CC   secretion), the detection, quantification and removal of histamine or
CC   serotonin (in animals, plants, cell cultures, food materials, or humans)
CC   and in the treatment of various diseases and allergies (e.g. type I
CC   hypersensitivity reactions, urticaria, asthma, allergic rhinitis (hay
CC   fever), atopic dermatitis, insect bites and food and drug allergies,
CC   abnormal blood pressure, migraine, psychological disorders, respiratory
CC   disease, and coronary heart disease). Histamine may also be used to
CC   regulate cellular growth and tissue repair. The molecules may also be
CC   used as components of vaccines directed against blood-sucking
CC   ectoparasites
XX
SQ   Sequence 190 AA;

  Query Match             25.8%;  Score 290.5;  DB 2;  Length 190;
  Best Local Similarity   36.0%;  Pred. No. 1.2e-20;
  Matches   72;  Conservative   32;  Mismatches   79;  Indels   17;  Gaps    9;

Qy          7 LLLTFVSAALATQAETTSAKAGENPLWAHEELLGKYQDAWKSIDQGVSVTYVLAKTTYEN 66
              ||:  ::  ||       |   |  | || |   | :||||||:   |   | : | ||:|
Db          3 LLILSLALVLA-----LSQVKGNQPDWADEAANGAHQDAWKSLKADVENVYYMVKATYKN 57

Qy         67 DTGSWGSQFKCLQVQEIERKEEDYTVTSVFTF-RNASSPIKYYNVTETVKAVFQYGYKNI 125
              |   ||: | |: |   :  |:: :: : | |  || : :::   || | ||  ||| |
Db         58 DP-VWGNDFTCVGVMANDVNEDEKSIQAEFLFMNNADTNMQF--ATEKVTAVKMYGY-NR 113

Qy        126 RNAIEYQVGGGLNITDTLIFTDGELCDVFYVPNAD---QGCELWVKKSHYKHVPDYCTFV 182
               ||  |:  |   || : ::| : ||| |||  |   :| |||   : | ::|  |
Db        114 ENAFRYETEDGQVFTDVIAYSD-DNCDVIYVPGTDGNEEGYELWT--TDYDNIPANCLNK 170

Qy        183 FNVFCAKDRKTYDIFNEECV 202
              || : |  |:| |:|   |:
Db        171 FNEY-AVGRETRDVFTSACL 189


RESULT 7
```

```
AAY18084
ID   AAY18084 standard; protein; 198 AA.
XX
AC   AAY18084;
XX
DT   06-AUG-1999  (first entry)
XX
DE   Histamine binding protein Av-HBP.
XX
KW   Histamine binding protein; serotonin binding compound; inflammation;
KW   gastric acid secretion; allergy; type I hypersensitivity reaction;
KW   asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;
KW   drug allergy; abnormal blood pressure; psychological disorder; vaccine;
KW   respiratory disease; coronary heart disease; cellular growth regulator;
KW   tissue repair; blood-sucking ectoparasite; therapy.
XX
OS   Amblyomma variegatum.
XX
PN   WO9927104-A1.
XX
PD   03-JUN-1999.
XX
PF   26-NOV-1998;   98WO-GB003530.
XX
PR   26-NOV-1997;   97GB-00025046.
PR   26-JUN-1998;   98GB-00013917.
XX
PA   (OXFO-) OXFORD VACS LTD.
XX
PI   Nuttall PA,  Paesen GC;
XX
DR   WPI; 1999-357841/30.
DR   N-PSDB; AAX76969.
XX
PT   Histamine and serotonin binding compounds useful for the treatment of
PT   allergies.
XX
PS   Claim 12; Fig 6; 84pp; English.
XX
CC   This sequence is an example of a histamine or serotonin binding compound
CC   (A), of the invention. The compounds are useful for regulating the action
CC   of histamine and serotonin (in e.g. inflammation and gastric acid
CC   secretion), the detection, quantification and removal of histamine or
CC   serotonin (in animals, plants, cell cultures, food materials, or humans)
CC   and in the treatment of various diseases and allergies (e.g. type I
CC   hypersensitivity reactions, urticaria, asthma, allergic rhinitis (hay
CC   fever), atopic dermatitis, insect bites and food and drug allergies,
CC   abnormal blood pressure, migraine, psychological disorders, respiratory
CC   disease, and coronary heart disease). Histamine may also be used to
CC   regulate cellular growth and tissue repair. The molecules may also be
CC   used as components of vaccines directed against blood-sucking
CC   ectoparasites
XX
SQ   Sequence 198 AA;

  Query Match             25.3%;  Score 284.5;  DB 2;  Length 198;
  Best Local Similarity   35.4%;  Pred. No. 5.2e-20;
  Matches   58;  Conservative   36;  Mismatches   63;  Indels    7;  Gaps    4;

Qy         31 PLWAHEELLGKYQDAWKSIDQGVSVTYVLAKTTYENDTGSWGSQFKCLQVQEIERKEEDY 90
              | || |    ||||:|||:::| :: |:|| ::|| ::    ||  | |:: : :|
Db         23 PTWADEGRFGKYQNAWKALNQRINTTHVLVRSTYIDNPYLWGKNFSCVRARTVEVFPSSK 82

Qy         91 TVTSVFTFRNASSPIKYYNVTETVKAVFQYGYKNIRNAIEYQVGGGLNITDTLIFTDGEL 150
              ||   |:||| :  :    |  :||:|    ||     || |: : |  : :: ::||||
Db         83 TVELEFSFRNRTGILCMRN--QTVRAGKDYFYHQ-PNAFEFMLRGNRSFSNAVMFTDGMT 139

Qy        151 CDVFYVP---NADQGCELWVKKSHYKHVPDYCTFVFNVFCAKDR 191
              |::   |   | | |||||| :   ::|  |:|:|:  | : |
Db        140 CNLLSFPYQRNKPQ-CELWVKDTRVDNIPPCCSFMFDYLCPQPR 182


RESULT 8
```

```
AAB73261
ID   AAB73261 standard; protein; 171 AA.
XX
AC   AAB73261;
XX
DT   15-MAY-2001  (first entry)
XX
DE   Histacalin protein FS-HBP2.
XX
KW   FS-HBP2; histacalin protein; antiinflammatory; antiallergic;
KW   opthalmological; allergic rhinitis.
XX
OS   Unidentified.
XX
PN   WO200116164-A2.
XX
PD   08-MAR-2001.
XX
PF   24-AUG-2000; 2000WO-GB003287.
XX
PR   01-SEP-1999;   99GB-00020673.
XX
PA   (EVOL-) EVOLUTEC LTD.
XX
PI   Nuttall PA,  Paesen GC;
XX
DR   WPI; 2001-218521/22.
XX
PT   Use of histacalin proteins for treating or preventing allergic rhinitis,
PT   or for manufacturing a medicament for treating or preventing allergic
PT   rhinitis, e.g. seasonal or perennial allergic rhinitis.
XX
PS   Disclosure; Page 4-6; 19pp; English.
XX
CC   The present invention relates to a method for treating or preventing
CC   allergic rhinitis. The method involves employing a blood-feeding
CC   ectoparasite-derived (e.g. tick-derived) histacalin protein. The present
CC   sequence is one such histacalin protein. The histacalin protein, is
CC   useful for treating or preventing allergic rhinitis, both seasonal and
CC   perennial allergic conjunctivitis
XX
SQ   Sequence 171 AA;

  Query Match             25.1%;  Score 282;  DB 4;  Length 171;
  Best Local Similarity   37.5%;  Pred. No. 7.6e-20;
  Matches   66;  Conservative   29;  Mismatches   69;  Indels   12;  Gaps    8;

Qy         31 PLWAHEELLGKYQDAWKSIDQGVSVTYVLAKTTYENDTGSWGSQFKCLQVQEIERKEEDY 90
              | || |    | :||||||:    |    | : | ||:||    ||: | |: |    :  |::
Db          3 PDWADEAANGAHQDAWKSLKADVENVYYMVKATYKNDP-VWGNDFTCVGVMANDVNEDEK 61

Qy         91 TVTSVFTF-RNASSPIKYYNVTETVKAVFQYGYKNIRNAIEYQVGGGLNITDTLIFTDGE 149
              :: : | |  || : :::   || | ||  ||| |  ||  |:   |   || : ::| :
Db         62 SIQAEFLFMNNADTNMQF--ATEKVTAVKMYGY-NRENAFRYETEDGQVFTDVIAYSD-D 117

Qy        150 LCDVFYVPNAD---QGCELWVKKSHYKHVPDYCTFVFNVFCAKDRKTYDIFNEECV 202
              ||| |||  |   :| |||   : | ::|  |   || : |  |:| |:|   |:
Db        118 NCDVIYVPGTDGNEEGYELWT--TDYDNIPANCLNKFNEY-AVGRETRDVFTSACL 170


RESULT 9
AAB74289
ID   AAB74289 standard; protein; 171 AA.
XX
AC   AAB74289;
XX
DT   20-JUN-2001  (first entry)
XX
DE   Histacalin protein FS-HBP2.
XX
KW   Histacalin; FS-HBP1; conjunctivitis.
XX
OS   Unidentified.
```

XX
PN WO200115719-A2.
XX
PD 08-MAR-2001.
XX
PF 24-AUG-2000; 2000WO-GB003282.
XX
PR 01-SEP-1999; 99GB-00020674.
XX
PA (EVOL-) EVOLUTEC LTD.
XX
PI Nuttall PA, Paesen GC;
XX
DR WPI; 2001-257675/26.
XX
PT Use of histacalin proteins for treating or preventing non-infective
PT conjunctivitis, or for manufacturing a medicament for treating or
PT preventing conjunctivitis, e.g. seasonal or perennial allergic
PT conjunctivitis.
XX
PS Claim 4; Page 5-6; 19pp; English.
XX
CC The present invention relates to the use of a histacalin protein for
CC treating or preventing conjunctivitis. The present sequence is the
CC histacalin protein FS-HBP1. The invention is particularly useful in the
CC treatment of allergic or seasonal conjunctivitis
XX
SQ Sequence 171 AA;

Query Match 25.1%; Score 282; DB 4; Length 171;
Best Local Similarity 37.5%; Pred. No. 7.6e-20;
Matches 66; Conservative 29; Mismatches 69; Indels 12; Gaps 8;

```
Qy         31 PLWAHEELLGKYQDAWKSIDQGVSVTYVLAKTTYENDTGSWGSQFKCLQVQEIERKEEDY 90
              | || |   | :||||||:   |   | : | ||:||   ||: | |: |   :  |::
Db          3 PDWADEAANGAHQDAWKSLKADVENVYYMVKATYKNDP-VWGNDFTCVGVMANDVNEDEK 61

Qy         91 TVTSVFTF-RNASSPIKYYNVTETVKAVFQYGYKNIRNAIEYQVGGGLNITDTLIFTDGE 149
              :: : | |  || : :::   || | ||  ||| |  ||  |:   |   || : ::| :
Db         62 SIQAEFLFMNNADTNMQF--ATEKVTAVKMYGY-NRENAFRYETEDGQVFTDVIAYSD-D 117

Qy        150 LCDVFYVPNAD---QGCELWVKKSHYKHVPDYCTFVFNVFCAKDRKTYDIFNEECV 202
               ||| |||  |   :| |||   : | ::|  |   || : |  |:| |:|   |:
Db        118 NCDVIYVPGTDGNEEGYELWT--TDYDNIPANCLNKFNEY-AVGRETRDVFTSACL 170
```

RESULT 10
AAW37448
ID AAW37448 standard; protein; 200 AA.
XX
AC AAW37448;
XX
DT 08-JUN-1998 (first entry)
XX
DE Tick vasoactive amine binding protein 1 MS-HBP1.
XX
KW Male-specific vasoactive amine binding protein 1; MS-HCP1; histamine;
KW serotonin; assay; antihistamine; anti-inflammatory; insect bite;
KW snake bite; scorpion bite; dermatitis; vaccine; transgenic animal; tick.
XX
OS Rhipicephalus appendiculatus.
XX
FH Key Location/Qualifiers
FT Peptide 1. .18
FT /label= Sig_peptide
FT Modified-site 79. .81
FT /note= "Asn is N-glycosylated"
XX
PN WO9744451-A2.
XX
PD 27-NOV-1997.
XX
PF 19-MAY-1997; 97WO-GB001372.
XX
PR 18-MAY-1996; 96GB-00010484.
PR 18-APR-1997; 97GB-00007844.
XX
PA (OXFO-) OXFORD VACS LTD.
XX
PI Paesen GC, Nuttall PA;
XX
DR WPI; 1998-018506/02.
DR N-PSDB; AAV00229.
XX
PT New vasoactive amine binding proteins and related nucleic acid, vectors -
PT transformed cells and transgenic animals, used for assaying or removing
PT histamine and as antihistamine or anti-inflammatory agents.
XX
PS Example 2; Fig 3; 44pp; English.
XX
CC This protein comprises tick Rhipicephalus appendiculatus (Ra) novel male-
CC specific histamine binding protein 1 (MS-HBP1). Its amino acid sequence
CC was deduced from a cDNA clone (see AAV00227) obtained from a salivary
CC gland cDNA library. Female-specific HSP1 and HSP2 (see AAW37446-47) and a
CC related protein, D.RET6 (see AAW37449) from Dermacenter reticularis, were
CC also identified. These novel vasoactive amine binding proteins (VABPs)
CC can be expressed in host cells using e.g. a baculovirus expression
CC system. They can be used: (i) to assay histamine (or other VA such as
CC serotonin) in body fluids or cell culture supernatants, e.g. to monitor
CC the effect of allergens; (ii) for binding VA, e.g. to remove histamine
CC from blood, food, cell cultures etc.; (iii) as an antihistamine or anti-
CC inflammatory agents, e.g. for treating insect, snake or scorpion bites or
CC dermatitis, or as a carrier for slow release of histamine-related
CC compounds; (iv) in vaccines to protect against metazoan parasites,
CC especially in animals; (v) as reagents for studying inflammation,
CC involvement of VA in ulcer formation or the immune response etc. VABPs
CC provide a more sensitive assay for histamine than low-affinity antibodies
CC currently used. They may also be more effective and safer than
CC conventional antihistamines
XX
SQ Sequence 200 AA;

Query Match 23.6%; Score 265; DB 2; Length 200;
Best Local Similarity 34.5%; Pred. No. 4.7e-18;
Matches 69; Conservative 34; Mismatches 71; Indels 26; Gaps 10;

```
Qy          3 MQVVLLLTFVSAALATQAETTSAKAGENPLWAHEELLGKYQDAWKSIDQGVSVTYVLAKT 62
              |:|:||:  : |||   |:          || ||:|  || |||||||: |  :  | ||:
Db          1 MKVLLLV--LGAALCQNADA-------NPTWANEAKLGSYQDAWKSLQQDQNKRYYLAQA 51

Qy         63 TYENDTGSWGSQFKCLQV--QEIERKEEDYTVTSVFTFRNAS-SPIKYYNVTETVKAVFQ 119
              |   | | || :| |: |  ::| :|: : |:     ::|   : :|  : | ||   :
Db         52 TQTTD-GVWGEEFTCVSVTAEKIGKKKLNATI----LYKNKHLTDLKESHETITVWKAYD 106

Qy        120 YGYKNIRNAIEYQVGGGLNIT--DTLIFTDGELCDVFYVP---NADQG-CELWVKKSHYK 173
              |       | |:|:  |    |  | :|:| : ||| :||    :|:|  |||| :
Db        107 Y---TTENGIKYETQGTRTQTFEDVFVFSDYKNCDVIFVPKERGSDEGDYELWVSEDKID 163

Qy        174 HVPDYCTFVFNVFCAKDRKT 193
              :|| | |    |  :  ||
Db        164 KIPDCCKFTMAYFAQQQEKT 183
```

RESULT 11
AAY18080
ID AAY18080 standard; protein; 200 AA.
XX
AC AAY18080;
XX
DT 06-AUG-1999 (first entry)
XX
DE Histamine binding protein MS-HBP1.
XX
KW Histamine binding protein; serotonin binding compound; inflammation;
KW gastric acid secretion; allergy; type I hypersensitivity reaction;
KW asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;

```
KW   drug allergy; abnormal blood pressure; psychological disorder; vaccine;
KW   respiratory disease; coronary heart disease; cellular growth regulator;
KW   tissue repair; blood-sucking ectoparasite; therapy.
XX
OS   Rhipicephalus appendiculatus.
XX
PN   WO9927104-A1.
XX
PD   03-JUN-1999.
XX
PF   26-NOV-1998;   98WO-GB003530.
XX
PR   26-NOV-1997;   97GB-00025046.
PR   26-JUN-1998;   98GB-00013917.
XX
PA   (OXFO-) OXFORD VACS LTD.
XX
PI   Nuttall PA,  Paesen GC;
XX
DR   WPI; 1999-357841/30.
DR   N-PSDB; AAX76966.
XX
PT   Histamine and serotonin binding compounds useful for the treatment of
PT   allergies.
XX
PS   Claim 1; Fig 3; 84pp; English.
XX
CC   This sequence is an example of a histamine or serotonin binding compound
CC   (A), of the invention. The compounds are useful for regulating the action
CC   of histamine and serotonin (in e.g. inflammation and gastric acid
CC   secretion), the detection, quantification and removal of histamine or
CC   serotonin (in animals, plants, cell cultures, food materials, or humans)
CC   and in the treatment of various diseases and allergies (e.g. type I
CC   hypersensitivity reactions, urticaria, asthma, allergic rhinitis (hay
CC   fever), atopic dermatitis, insect bites and food and drug allergies,
CC   abnormal blood pressure, migraine, psychological disorders, respiratory
CC   disease, and coronary heart disease). Histamine may also be used to
CC   regulate cellular growth and tissue repair. The molecules may also be
CC   used as components of vaccines directed against blood-sucking
CC   ectoparasites
XX
SQ   Sequence 200 AA;

  Query Match             23.6%;  Score 265;  DB 2;  Length 200;
  Best Local Similarity   34.5%;  Pred. No. 4.7e-18;
  Matches   69;  Conservative   34;  Mismatches   71;  Indels   26;  Gaps   10;

Qy          3 MQVVLLLTFVSAALATQAETTSAKAGENPLWAHEELLGKYQDAWKSIDQGVSVTYVLAKT 62
              |:|:||:  : |||   |:        || ||:|  || |||||||: |  :  | ||:
Db          1 MKVLLLV--LGAALCQNADA-------NPTWANEAKLGSYQDAWKSLQQDQNKRYYLAQA 51

Qy         63 TYENDTGSWGSQFKCLQV--QEIERKEEDYTVTSVFTFRNAS-SPIKYYNVTETVKAVFQ 119
              |   | | || :| |: |  ::| :|: : |:     ::|   : :|  : | ||   :
Db         52 TQTTD-GVWGEEPTCVSVTAEKIGKKKLNATI----LYKNKHLTDLKESHETITVWKAYD 106

Qy        120 YGYKNIRNAIEYQVGGGLNIT--DTLIFTDGELCDVFYVP---NADQG-CELWVKKSHYK 173
              |       | |:|:  |    |  |  :|:| : ||| :||    :|:|  |||| :
Db        107 Y---TTENGIKYETQGTRTQTFEDVFVFSDYKNCDVIFVPKERGSDEGDYELWVSEDKID 163

Qy        174 HVPDYCTFVFNVFCAKDRKT 193
              :|| | |     |  :  ||
Db        164 KIPDCCKFTMAYFAQQQEKT 183


RESULT 12
AAW37446
ID   AAW37446 standard; protein; 190 AA.
XX
AC   AAW37446;
XX
DT   08-JUN-1998  (first entry)
XX
DE   Tick vasoactive amine binding protein 1 FS-HBP1.
XX
KW   Female-specific vasoactive amine binding protein 1; FS-HCP1; histamine;
KW   serotonin; assay; antihistamine; anti-inflammatory; insect bite;
KW   snake bite; scorpion bite; dermatitis; vaccine; transgenic animal; tick.
XX
OS   Rhipicephalus appendiculatus.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .18
FT                   /label= Sig_peptide
XX
PN   WO9744451-A2.
XX
PD   27-NOV-1997.
XX
PF   19-MAY-1997;   97WO-GB001372.
XX
PR   18-MAY-1996;   96GB-00010484.
PR   18-APR-1997;   97GB-00007844.
XX
PA   (OXFO-) OXFORD VACS LTD.
XX
PI   Paesen GC,  Nuttall PA;
XX
DR   WPI; 1998-018506/02.
DR   N-PSDB; AAV00227.
XX
PT   New vasoactive amine binding proteins and related nucleic acid, vectors -
PT   transformed cells and transgenic animals, used for assaying or removing
PT   histamine and as antihistamine or anti-inflammatory agents.
XX
PS   Example 2; Fig 1; 44pp; English.
XX
CC   This protein comprises tick Rhipicephalus appendiculatus (Ra) novel
CC   female-specific histamine binding protein 1 (FS-HBP1). Its amino acid
CC   sequence was deduced from a cDNA clone (see AAV00227) obtained from a
CC   salivary gland cDNA library. FS-HSP2 and male-specific HSP1 (see AAW37447
CC   -48) and a related protein, D.RET6 (see AAW37449) from Dermacenter
CC   reticularis, were also identified. These novel vasoactive amine binding
CC   proteins (VABPs) can be expressed in host cells using e.g. a baculovirus
CC   expression system. They can be used: (i) to assay histamine (or other VA
CC   such as serotonin) in body fluids or cell culture supernatants, e.g. to
CC   monitor the effect of allergens; (ii) for binding VA, e.g. to remove
CC   histamine from blood, food, cell cultures etc.; (iii) as an antihistamine
CC   or anti-inflammatory agents, e.g. for treating insect, snake or scorpion
CC   bites or dermatitis, or as a carrier for slow release of histamine-
CC   related compounds; (iv) in vaccines to protect against metazoan
CC   parasites, especially in animals; (v) as reagents for studying
CC   inflammation, involvement of VA in ulcer formation or the immune response
CC   etc. VABPs provide a more sensitive assay for histamine than low-affinity
CC   antibodies currently used. They may also be more effective and safer than
CC   conventional antihistamines
XX
SQ   Sequence 190 AA;

  Query Match             22.4%;  Score 252;  DB 2;  Length 190;
  Best Local Similarity   33.5%;  Pred. No. 8.7e-17;
  Matches   68;  Conservative   28;  Mismatches   89;  Indels   18;  Gaps    8;

Qy          3 MQVVLLLTFVSAALATQAETTSAKAGENPLWAHEELLGKYQDAWKSIDQGVSVTYVLAKT 62
              |:::| | || |    :|        : |:|| |   |::||||| : : |   | | |
Db          1 MKLLLSLAFVLALSQVKA--------DKPVWADEAANGEHQDAWKHLQKLVEENYDLIKA 52

Qy         63 TYENDTGSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASSPIKYYNVTETVKAVFQYGY 122
              ||:||   ||: | |:         |::  | : | | | :  : | : |   |||
Db         53 TYKNDP-VWGNDFTCVGTAAQNLNEDEKNVEAWFMFMNNADTV-YQHTFEKATPDKMYGY 110

Qy        123 KNIRNAIEYQVGGGLNITDTLIFTDGELCDVFYVPNAD---QGCELWVKKSHYKHVPDYC 179
              |  ||| ||    | :|| | |:| : | | |    |   | |||   : |  ||  |
Db        111 -NKENAITYQTEDGQVLTDVLAFSD-DNCYVIYALGPDGSGAGYELWA--TDYTDVPASC 166

Qy        180 TFVFNVFCAKDRKTYDIFNEECV 202
                || : |      |::  :|:
```

```
Db        167 LEKFNEYAA-GLPVRDVYTSDCL 188


RESULT 13
AAY18078
ID   AAY18078 standard; protein; 190 AA.
XX
AC   AAY18078;
XX
DT   06-AUG-1999  (first entry)
XX
DE   Histamine binding protein FS-HBP1.
XX
KW   Histamine binding protein; serotonin binding compound; inflammation;
KW   gastric acid secretion; allergy; type I hypersensitivity reaction;
KW   asthma; urticaria; allergic rhinitis; atopic dermatitis; food allergy;
KW   drug allergy; abnormal blood pressure; psychological disorder; vaccine;
KW   respiratory disease; coronary heart disease; cellular growth regulator;
KW   tissue repair; blood-sucking ectoparasite; therapy.
XX
OS   Rhipicephalus appendiculatus.
XX
PN   WO9927104-A1.
XX
PD   03-JUN-1999.
XX
PF   26-NOV-1998;   98WO-GB003530.
XX
PR   26-NOV-1997;   97GB-00025046.
PR   26-JUN-1998;   98GB-00013917.
XX
PA   (OXFO-) OXFORD VACS LTD.
XX
PI   Nuttall PA,  Paesen GC;
XX
DR   WPI; 1999-357841/30.
DR   N-PSDB; AAX76964.
XX
PT   Histamine and serotonin binding compounds useful for the treatment of
PT   allergies.
XX
PS   Claim 1; Fig 1; 84pp; English.
XX
CC   This sequence is an example of a histamine or serotonin binding compound
CC   (A), of the invention. The compounds are useful for regulating the action
CC   of histamine and serotonin (in e.g. inflammation and gastric acid
CC   secretion), the detection, quantification and removal of histamine or
CC   serotonin (in animals, plants, cell cultures, food materials, or humans)
CC   and in the treatment of various diseases and allergies (e.g. type I
CC   hypersensitivity reactions, urticaria, asthma, allergic rhinitis (hay
CC   fever), atopic dermatitis, insect bites and food and drug allergies,
CC   abnormal blood pressure, migraine, psychological disorders, respiratory
CC   disease, and coronary heart disease). Histamine may also be used to
CC   regulate cellular growth and tissue repair. The molecules may also be
CC   used as components of vaccines directed against blood-sucking
CC   ectoparasites
XX
SQ   Sequence 190 AA;

  Query Match             22.4%;  Score 252;  DB 2;  Length 190;
  Best Local Similarity   33.5%;  Pred. No. 8.7e-17;
  Matches   68;  Conservative   28;  Mismatches   89;  Indels   18;  Gaps    8;

Qy          3 MQVVLLLTFVSAALATQAETTSAKAGENPLWAHEELLGKYQDAWKSIDQGVSVTYVLAKT 62
              |:::| | || |    :|       : |:|| |   |::||||| : : |   | | |
Db          1 MKLLLSLAFVLALSQVKA--------DKPVWADEAANGEHQDAWKHLQKLVEENYDLIKA 52

Qy         63 TYENDTGSWGSQFKCLQVQEIERKEEDYTVTSVFTFRNASSPIKYYNVTETVKAVFQYGY 122
              ||:||   ||: | |:          |::  | : | | :  : | :  |         |||
Db         53 TYKNDP-VWGNDFTCVGTAAQNLNEDEKNVEAWFMFMNNADTV-YQHTFEKATPDKMYGY 110

Qy        123 KNIRNAIEYQVGGGLNITDTLIFTDGELCDVFYVPNAD---QGCELWVKKSHYKHVPDYC 179
                |  ||| ||   |  :|| | |:| : | | |    |    | |||   : |  ||  |

Db        111 -NKENAITYQTEDGQVLTDVLAFSD-DNCYVIYALGPDGSGAGYELWA--TDYTDVPASC 166

Qy        180 TFVFNVFCAKDRKTYDIFNEECV 202
                || : |       |::  :|:
Db        167 LEKFNEYAA-GLPVRDVYTSDCL 188


RESULT 14
AAB73262
ID   AAB73262 standard; protein; 182 AA.
XX
AC   AAB73262;
XX
DT   15-MAY-2001  (first entry)
XX
DE   Histacalin protein MS-HBP1.
XX
KW   MS-HBP1; histacalin protein; antiinflammatory; antiallergic;
KW   opthalmological; allergic rhinitis.
XX
OS   Unidentified.
XX
PN   WO200116164-A2.
XX
PD   08-MAR-2001.
XX
PF   24-AUG-2000; 2000WO-GB003287.
XX
PR   01-SEP-1999;   99GB-00020673.
XX
PA   (EVOL-) EVOLUTEC LTD.
XX
PI   Nuttall PA,  Paesen GC;
XX
DR   WPI; 2001-218521/22.
XX
PT   Use of histacalin proteins for treating or preventing allergic rhinitis,
PT   or for manufacturing a medicament for treating or preventing allergic
PT   rhinitis, e.g. seasonal or perennial allergic rhinitis.
XX
PS   Disclosure; Page 4-6; 19pp; English.
XX
CC   The present invention relates to a method for treating or preventing
CC   allergic rhinitis. The method involves employing a blood-feeding
CC   ectoparasite-derived (e.g. tick-derived) histacalin protein. The present
CC   sequence is one such histacalin protein. The histacalin protein, is
CC   useful for treating or preventing allergic rhinitis, both seasonal and
CC   perennial allergic conjunctivitis
XX
SQ   Sequence 182 AA;

  Query Match             22.4%;  Score 251.5;  DB 4;  Length 182;
  Best Local Similarity   35.3%;  Pred. No. 9.2e-17;
  Matches   61;  Conservative   29;  Mismatches   66;  Indels   17;  Gaps    8;

Qy         30 NPLWAHEELLGKYQDAWKSIDQGVSVTYVLAKTTYENDTGSWGSQFKCLQV--QEIERKE 87
              || ||:|  || |||||||: |  :  | ||: |   | | || :| |: |  ::| :|:
Db          1 NPTWANEAKLGSYQDAWKSLQQDQNKRYYLAQATQTTD-GVWGEEFTCVSVTAEKIGKKK 59

Qy         88 EDYTVTSVFTFRNAS-SPIKYYNVTETVKAVFQYGYKNIRNAIEYQVGGGLNIT--DTLI 144
                : |:     ::|  : :| : | ||   : |       | |:|:  |     |  |  :
Db         60 LNATI----LYKNKHLTDLKESHETITVWKAYDY---TTENGIKYETQGTRTQTFEDVFV 112

Qy        145 FTDGELCDVFYVP---NADQG-CELWVKKSHYKHVPDYCTFVFNVFCAKDRKT 193
              |:| : ||| :||    :|:|  |||| :      :|| | |    |  :  ||
Db        113 FSDYKNCDVIFVPKERGSDEGDYELWVSEDKIDKIPDCCKFTMAYFAQQQEKT 165


RESULT 15
AAB74290
ID   AAB74290 standard; protein; 182 AA.
XX
AC   AAB74290;
```

XX
DT 20-JUN-2001 (first entry)
XX
DE Histacalin protein MS-HBP1.
XX
KW Histacalin; MS-HBP1; conjunctivitis.
XX
OS Unidentified.
XX
PN WO200115719-A2.
XX
PD 08-MAR-2001.
XX
PF 24-AUG-2000; 2000WO-GB003282.
XX
PR 01-SEP-1999; 99GB-00020674.
XX
PA (EVOL-) EVOLUTEC LTD.
XX
PI Nuttall PA, Paesen GC;
XX
DR WPI; 2001-257675/26.
XX
PT Use of histacalin proteins for treating or preventing non-infective
PT conjunctivitis, or for manufacturing a medicament for treating or
PT preventing conjunctivitis, e.g. seasonal or perennial allergic
PT conjunctivitis.
XX
PS Claim 4; Page 5-6; 19pp; English.
XX
CC The present invention relates to the use of a histacalin protein for
CC treating or preventing conjunctivitis. The present sequence is the
CC histacalin protein FS-HBP1. The invention is particularly useful in the
CC treatment of allergic or seasonal conjunctivitis
XX
SQ Sequence 182 AA;

Query Match 22.4%; Score 251.5; DB 4; Length 182;
Best Local Similarity 35.3%; Pred. No. 9.2e-17;
Matches 61; Conservative 29; Mismatches 66; Indels 17; Gaps 8;

```
Qy          30 NPLWAHEELLGKYQDAWKSIDQGVSVTYVLAKTTYENDTGSWGSQFKCLQV--QEIERKE 87
               || ||:|  || |||||||: |  :  | ||: |   | | || :| |: |  ::| :|:
Db           1 NPTWANEAKLGSYQDAWKSLQQDQNKRYYLAQATQTTD-GVWGEEFTCVSVTAEKIGKKK 59

Qy          88 EDYTVTSVFTFRNAS-SPIKYYNVTETVKAVFQYGYKNIRNAIEYQVGGGLNIT--DTLI 144
                : |:     ::|   : :|  : | ||   : |       | |:|:  |    |  |  :
Db          60 LNATI----LYKNKHLTDLKESHETITVWKAYDY---TTENGIKYETQGTRTQTFEDVFV 112

Qy         145 FTDGELCDVFYVP---NADQG-CELWVKKSHYKHVPDYCTFVFNVFCAKDRKT 193
               |:| : ||| :||    :|:|  |||| :     :|| | |     |  :  ||
Db         113 FSDYKNCDVIFVPKERGSDEGDYELWVSEDKIDKIPDCCKFTMAYFAQQQEKT 165
```

Search completed: March 9, 2004, 13:33:28
Job time : 62 secs